



UCA

Universidad
de Cádiz

Grado en Biotecnología
Facultad de Ciencias
Curso 2020/21

Caracterización estructural de genes de determinación sexual en el lenguado senegalés

Pilar González Marchante

Tutores:

- Alberto Arias Pérez
- Laureana Rebordinos
González



Facultad de Ciencias
<http://ciencias.uca.es>

Biotecnología
Enología
Ingeniería Química
Matemáticas
Química

ÍNDICE

1. Abreviaturas.....	1
2. Resumen.....	2
3. Abstract.....	3
4. Introducción.....	4
4.1. Biología del lenguado senegalés.....	4
4.2. Cultivo del lenguado senegalés.....	5
4.3. Mecanismos de determinación sexual en vertebrados.....	8
4.4. Gen <i>dmrt1</i> en el lenguado senegalés y otras especies relacionadas.....	10
5. Objetivos.....	12
6. Materiales y métodos.....	13
6.1. Diseño de <i>primers</i>	14
6.2. PCR <i>in silico</i>	17
6.3. Extracción de ADN genómico.....	17
6.4. Amplificación por PCR.....	18
6.5. Electroforesis en gel de agarosa.....	20
6.6. Purificación de producto PCR-10 y escisión y reamplificación de bandas.....	21
6.7. Digestión con enzimas de restricción del intrón 3.....	23
6.8. Retrotranscripción de ARN a ADNc.....	24
6.9. PCR cuantitativa (qPCR).....	25
6.10. Alineamiento de la secuencia del gen <i>dmrt1</i> con el genoma.....	26
7. Resultados y discusión.....	27
7.1. Diseño de <i>primers</i>	27
7.2. PCR <i>in silico</i>	30
7.3. Extracción de ADN genómico.....	32
7.4. Amplificación por PCR de fragmentos largos (con <i>primers dmrt1</i> -10, 11, 15 y 16).....	32
7.5. Amplificación por PCR del intrón 3 (con <i>primers dmrt1</i> -12).....	34
7.6. Amplificación por PCR de fragmentos cortos (con <i>primers dmrt1</i> -13 y 14).....	36
7.7. Escisión y reamplificación de banda de 1.800 pb de PCR-10.....	39
7.8. Análisis de polimorfismos mediante digestión con enzimas de restricción del intrón 3.....	40
7.9. Retrotranscripción de ARN a ADNc.....	42
7.10. PCR cuantitativa (qPCR) (con <i>primers dmrt1</i> -13 y 14).....	42
7.11. Alineamiento de la secuencia del gen <i>dmrt1</i> con el genoma.....	45
8. Conclusiones.....	47
9. Perspectivas de futuro.....	48
10. Bibliografía.....	49
11. Agradecimientos.....	51
12. Anexos.....	51
12.1. Alineamiento de <i>dmrt1</i> con el genoma por BLAST+ (versión local).....	51
12.2. Alineamiento de <i>dmrt1</i> con el genoma por blastn (versión online).....	51

PROPUESTA DE TFG ASIGNADA

Titulación: GBT

Título: Caracterización estructural de genes de determinación sexual en el lenguado senegalés (nueva propuesta)

Curso: 2020/2021

Tutor1: alberto.arias@uca.es (BIOMEDICINA, BIOTECNOLOGÍA Y SALUD PÚBLICA)

Tutor2: laureana.rebordinos@uca.es (BIOMEDICINA, BIOTECNOLOGÍA Y SALUD PÚBLICA)

Carácter: Trabajos de iniciación a la investigación

Descripción:

El lenguado senegalés *Solea senegalensis* (Kaup, 1858) es una especie de gran interés económico en el sur de Europa, principalmente en las costas del oeste y sur de España. Sin embargo, el desarrollo de su acuicultura se ha visto dificultado por problemas relacionados con la reproducción en cautividad, crecimiento y mortalidad de las larvas y prevención de patologías. Un mayor conocimiento del genoma del lenguado ayudaría a comprender distintos aspectos de la biología de esta especie y facilitaría el desarrollo y mejora de su cultivo.

En los procesos de determinación y diferenciación sexual de vertebrados e invertebrados se ha observado la participación de distintos genes. Sin embargo, para el lenguado estos procesos no se conocen adecuadamente. El estudio de la estructura génica de genes potencialmente implicados facilitaría la comprensión de estos procesos.

En este estudio se pretende analizar la estructura genómica de genes (p. ej. *dmrt1*) implicados en la determinación sexual del lenguado. Para ello, se partirá de ADN genómico de la especie y se amplificará utilizando cebadores diseñados en base a trabajos previos o datos de otras especies. Posteriormente se secuenciarán y analizarán las lecturas de ADN obtenidas.

Alumno propuesto: Pilar González Marchante pi.gonmar@alum.uca.es

Curso: 2020/2021

Fecha: 06/11/2020 10:39:43

Es propuesta de alumno: NO; **O Aceptada profesor:** SÍ; **Aceptada depto.:** SÍ; **Aceptada comisión:** SÍ.

1. ABREVIATURAS

- **ADNc**: ADN complementario.
- **ADNg**: ADN genómico.
- **ARNm**: ARN mensajero.
- **Dmrt1**: *doublesex and male abnormal-3 related transcription factor 1*.
- **E-value**: valor esperado o valor E (*expected value*).
- **F1**: primera generación o filial 1, cultivado (*first filial generation*).
- **Fw**: con sentido (*forward*).
- **GWAS**: estudios de asociación del genoma completo (*genome-wide association study*).
- **LG**: grupo de ligamiento, es decir, todos los genes que pertenecen a un mismo cromosoma (*linkage group*).
- **ORF**: marco de lectura abierto (*open-reading frame*).
- **pb**: pares de bases.
- **PCR**: reacción en cadena de la polimerasa (*polymerase chain reaction*).
- **qPCR**: PCR cuantitativa.
- **QTL**: loci de caracteres cuantitativos (*quantitative trait loci*).
- **Rev**: antisentido (*reverse*).
- **Rxn**: reacción.
- **SDR**: región determinante del sexo (*sex-determining region*).
- **SNP**: polimorfismo de un solo nucleótido (*single-nucleotide polymorphism*).
- **Tm**: temperatura de fusión (*melting temperature*).
- **WT**: salvaje (*wild type*).

2. RESUMEN

Solea senegalensis es un pez plano que fue introducido en el sector de la acuicultura del suroeste de Europa durante la década de los 80 en un intento de diversificar el mercado. Aunque se han hecho muchos esfuerzos para impulsar su cultivo que han llevado a un aumento exponencial de su producción en los últimos años, hoy en día este aún se encuentra limitado por una serie de problemas entre los que destaca la complicada reproducción en cautividad debida a una ausencia de cortejo sexual en los machos cultivados, a lo que se le suma la dificultad de distinguir machos de hembras.

Si se consiguiera dilucidar el mecanismo de determinación sexual para esta especie, todavía hoy desconocido, se podría poner fin a muchas de estas cuestiones que dificultan su cultivo, además de dar paso a nuevas técnicas que lo mejoren, como la obtención de poblaciones exclusivamente femeninas, que por su biología alcanzan mayores tamaños y crecen más rápidamente que los machos. En los peces teleósteos este mecanismo puede estar definido por componentes genéticos y ambientales, a menudo dándose una mezcla de ambos. Si bien no existe un gen maestro regulador universal como en mamíferos, se han hallado genes como *dmrt1* (*doublesex and male abnormal-3 related transcription factor 1*), un factor de transcripción muy conservado en la evolución que presenta un importante papel en la regulación del desarrollo masculino en los vertebrados. Este gen autosómico se encuentra duplicado en el cromosoma Y de *Oryzias latipes* conformando el gen *dmy*, que se trata del regulador maestro en esta especie.

Recientemente se ha descrito por primera vez la estructura de *dmrt1* en el lenguado senegalés, que con unas 31.400 pb está compuesto de 7 exones y 6 intrones, en el que se observa una duplicación interna que abarca un tercio del gen. Asimismo, a partir de las secuencias de ADN complementario se han reportado 6 transcritos de ARN mensajero diferentes, y aunque no se ha encontrado una relación con el sexo de los individuos, podrían estar relacionados con la diferenciación sexual.

Por todo ello, en el presente trabajo se trató de obtener y caracterizar la secuencia completa del gen para distintos individuos, con el objetivo de buscar diferencias que confirmaran el papel de *dmrt1* en la determinación y diferenciación sexual de *S. senegalensis*. También se quiso examinar la diversidad del intrón 3 mediante digestiones con enzimas de restricción en búsqueda de polimorfismos que pudieran ser sexo-específicos. Por último, se planteó un estudio de expresión génica de *dmrt1* por PCR cuantitativa para múltiples individuos.

Debido a problemas durante las amplificaciones largas por PCR, no fue posible secuenciar los individuos, pero sí que se pudo confirmar la secuencia previamente descrita con un genoma del lenguado recién publicado. Por otro lado, no se encontraron polimorfismos en la secuencia del intrón 3 de *dmrt1* con las enzimas de restricción empleadas. Finalmente, no se logró cuantificar la expresión del gen por problemas de amplificación desconocidos al emplear ADN complementario y por una formación excesiva de dímeros de *primers*.

3. ABSTRACT

Solea senegalensis is a flatfish that was introduced into the aquaculture sector in southwestern Europe during the 1980s in an attempt to diversify the market. Although many efforts have been made to promote its cultivation, which have led to an exponential increase in its production in recent years, today it is still limited by a series of problems, among which is the complicated reproduction in captivity due to the absence of sexual courtship in cultured males, in addition to the difficulty of distinguishing males from females.

If the mechanism of sexual determination for this species, still unknown today, could be elucidated, it could put an end to many of these issues that plague its culture, in addition to giving way to new techniques that could improve it, such as obtaining exclusively female populations, which due to their biology reach larger sizes and grow faster than males. In teleost fish, this mechanism may be defined by genetic and environmental components, often a mixture of both. Although there is no universal master regulator gene as there is in mammals, genes such as *dmrt1* (*doublesex and male abnormal-3 related transcription factor 1*) have been found. This is a transcription factor that has been highly conserved throughout evolution and plays an important role in the regulation of male development in vertebrates. This autosomal gene is duplicated on the Y chromosome of *Oryzias latipes* forming the *dmy* gene, which is the master regulator in this species.

Recently, the structure of *dmrt1* has been described for the first time in the senegalese sole, which is composed of 7 exons and 6 introns with a total length of approximately 31,400 bp and an internal duplication that covers one third of the gene. Likewise, 6 different messenger RNA transcripts have been reported from the complementary DNA sequences, and although no relationship with the sex of the individuals has been found, they could be related to sexual regulation.

For all these reasons, in the present work we tried to obtain and characterize the complete sequence of this gene for different individuals, with the aim of searching for differences that would confirm the role of *dmrt1* in sex determination and differentiation in *S. senegalensis*. We also wanted to examine the diversity of intron 3 through restriction enzyme digests in search of polymorphisms that could, potentially, be sex-specific. Finally, a gene expression study of *dmrt1* by quantitative PCR was proposed for multiple individuals.

Due to problems during long PCR amplification, it was not possible to sequence the individuals, but we were able to confirm the previously described sequence against a recently published sole genome. On the other hand, no polymorphisms were found in the sequence of intron 3 of *dmrt1* with the restriction enzymes that were employed. Finally, gene expression could not be quantified due to unknown amplification problems when using complementary DNA and excessive primer dimer formation.

4. INTRODUCCIÓN

4.1. Biología del lenguado senegalés

El lenguado senegalés (*Solea senegalensis*) es una especie de pez plano, con cuerpo ovalado, alargado y asimétrico con los ojos en el lado derecho del cuerpo, incluida en el orden de los Pleuronectiformes y dentro de la familia Soleidae [1, 2]. Este teleósteo está estrechamente relacionado con el lenguado común (*Solea solea*) con el que comparte un alto nivel de potencial dentro del sector de la acuicultura, si bien este se encuentra mejor valorado en el centro y norte de Europa, mientras que el lenguado senegalés presenta mayor interés comercial en el sur del continente debido a una mayor tasa de crecimiento y unos requerimientos de temperatura más altos [3].

La cara ocular presenta una coloración muy variable puesto que es capaz de adaptarse miméticamente al sustrato principal del lecho marino (dado que, como todo pez plano, presenta un estilo de vida bentónico), aunque suele ir de un marrón verdoso a uno marrón rojizo, con ocelos azules distribuidos a lo largo de todo el cuerpo como se observa en la figura 1; la cara ciega, por su parte, es completamente blanca [4, 5]. La aleta pectoral de la cara ocular tiene una membrana interradianal negra con radios amarillos, carácter que lo diferencia del lenguado común, cuya aleta presenta en su lugar una mancha negra en su mitad posterior. En el Mediterráneo llega a alcanzar los 45 cm de longitud e incluso los 60 cm en el océano Atlántico, con pesos alrededor de los 3 kg, siendo las hembras mayores que los machos [5].

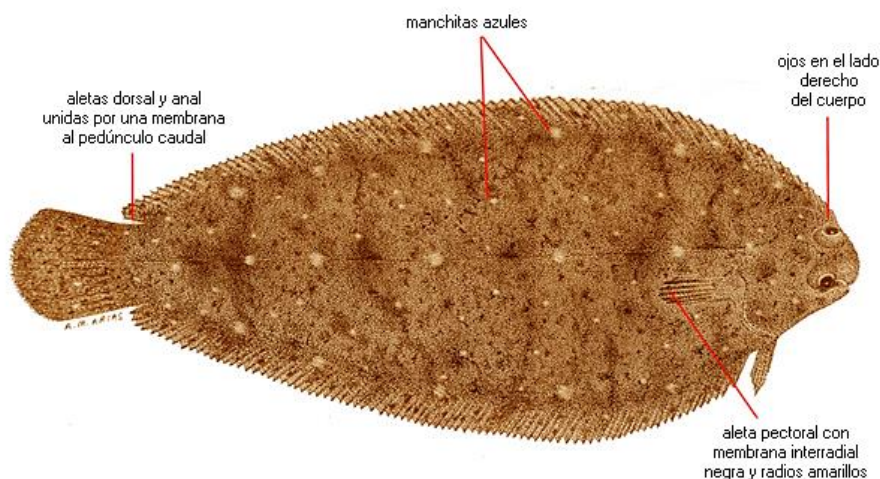


Figura 1. Rasgos característicos más importantes de *S. senegalensis*. Fuente: A. M. Arias. Recuperado de http://www.ictieterm.es/nombre_cientifico.php?nc=183.

Se trata de una especie dioica (también llamada gonocórica, es decir, con sexos separados en todas las etapas de su vida) sin caracteres sexuales externos diferenciados, que en promedio alcanza la madurez sexual a los 3 años de edad con un tamaño de 30 cm. Las hembras crecen más rápido y maduran más tarde que los machos [6]. Pese a la ausencia de estos caracteres, en época de puesta las hembras son fáciles de identificar por una considerable dilatación del ovario maduro, que recorre al pez por completo de la cabeza a la cola [1, 2, 5, 6].

En el medio natural es capaz de reproducirse prácticamente todo el año, si bien la puesta de huevos presenta un pico máximo en los meses de marzo a junio. En la naturaleza el tiempo de incubación ronda una semana, aunque depende de la temperatura del agua [1, 2]. Tras la eclosión, las larvas son inicialmente simétricas y después de unos 10-12 días de vida experimentan una metamorfosis, que dura una semana, en la que se transforman en juveniles asimétricos y con hábito de vida bentónico, al haberse desplazado el ojo izquierdo hacia el costado derecho, que a partir de ahora será la cara ocular. Este proceso de migración ocular es común a todos los peces planos, pero en el caso de la familia Soleidae la metamorfosis se inicia a una de las edades más tempranas [1, 2].

A partir de este momento, habita sobre sustratos arenosos desde el entorno intermareal, en lagunas salobres, hasta profundidades de 100 m en el mar. Durante el resto de su ciclo de vida, permanece durante el día semienterrado (al ser un mal nadador con hábito de vida nocturna) y se alimenta a base de pequeños invertebrados bentónicos (larvas de poliquetos y moluscos bivalvos) y en menor medida, de pequeños crustáceos. Con respecto a su distribución, este lenguado se encuentra principalmente en el Atlántico del este, desde el Golfo de Vizcaya hasta las costas de Senegal; y menos frecuentemente en las aguas del Mediterráneo Occidental y en el norte de África, tal y como se ve en la figura 2 [1, 2].

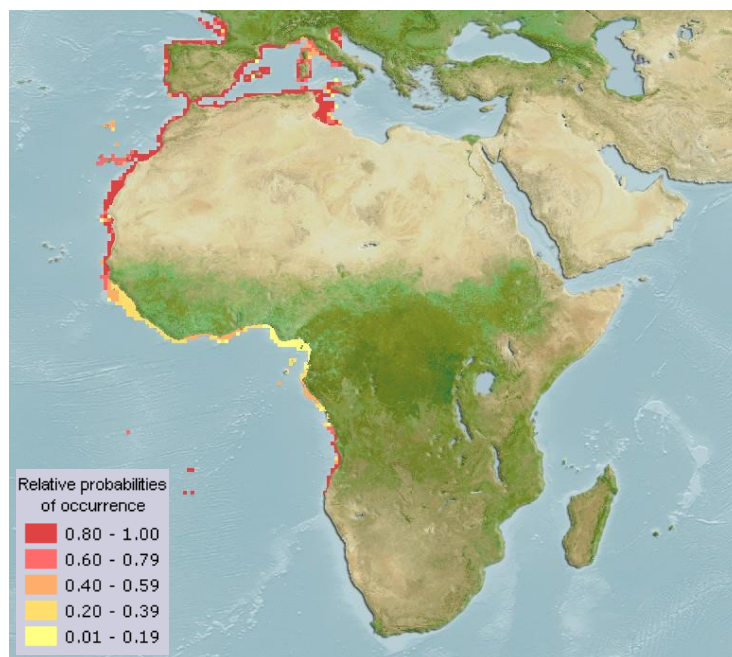


Figura 2. Mapa de distribución nativa actual de *S. senegalensis*. Fuente: AquaMaps (2019, octubre). Recuperado el 26 de mayo de 2021 de <http://www.aquamaps.org>.

4.2. Cultivo del lenguado senegalés

El cultivo del lenguado, en general, comenzó en Portugal y España a principios de los años 80 como un intento de diversificar la acuicultura del sur de Europa, que en aquel momento se encontraba saturada por la dorada, la lubina y el rodaballo. Concretamente se escogió al lenguado senegalés como candidato más prometedor frente al lenguado común debido a que ya se venía cultivando tradicionalmente en esta zona [7], además de por los hábitos alimenticios de la población de estos países, su precio y la elevada demanda en el mercado [2].

Desde entonces, a pesar de los grandes esfuerzos que se han llevado a cabo para implementar su cultivo a gran escala, tales como la obtención de puestas naturales de forma continuada en el centro IFAPA El Toruño (Cádiz) para su distribución a empresas y centros de investigación, y un inmenso volumen de investigaciones sobre el cultivo de la especie, todavía existe una serie de desafíos que deben ser superados para que el lenguado senegalés logre consolidarse a nivel de producción industrial. Estos problemas, entre los que se incluye la dependencia de ejemplares capturados en el medio natural, serán desarrollados con mayor detalle posteriormente [2, 8].

Pese a estas complicaciones, su cultivo ha aumentado exponencialmente en años recientes: en 2019 la cosecha mundial sobrepasó las 1.650 toneladas, siendo los mayores productores España, Islandia, Francia y Portugal, en ese orden [9]. En ese mismo año, la producción de lenguado de acuicultura en España fue de 818 toneladas (figura 3), localizándose principalmente en Galicia y Andalucía. A pesar de esta tendencia general positiva de las últimas décadas, se estima una disminución del 27% para la cosecha en España en el año 2020, bajando hasta las 597 toneladas, lo cual se atribuye a la pandemia de COVID-19, que ha afectado gravemente al sector de la acuicultura [9].

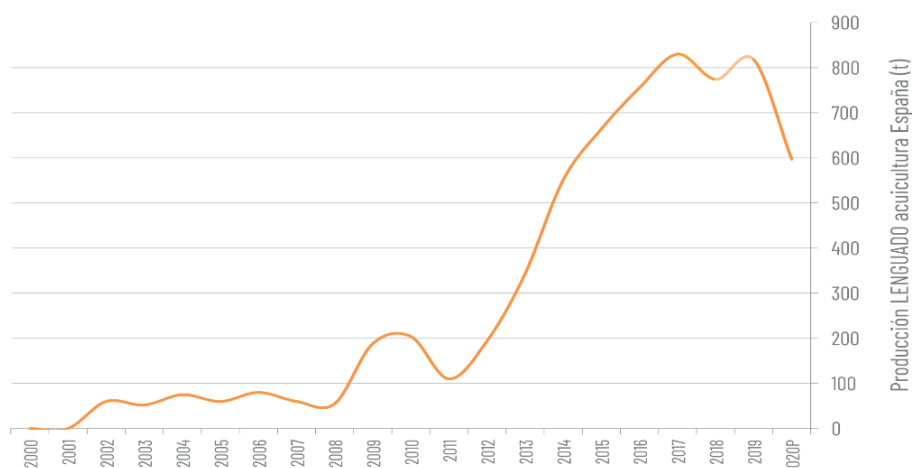


Figura 3. Evolución de la cosecha (producción de acuicultura) del lenguado (*S. senegalensis*) en España en toneladas (2000-2020). Fuente: informe de APROMAR “La Acuicultura en España 2020” [9].

Entre los retos que la acuicultura del lenguado senegalés todavía debe superar destaca principalmente la dificultad de conseguir puestas de huevos regulares de manera espontánea a partir de reproductores mantenidos en cautividad procedentes de primera generación (F1), que podría estar explicada en parte por la ausencia de cortejo sexual por parte de los individuos F1, como se ha observado en estudios por grabación [1, 10].

Si bien no se conoce por qué existe este fallo de cortejo sexual, hay algunas hipótesis que señalan la posible influencia de factores externos durante el desarrollo temprano del lenguado que podrían condicionar la facultad para el cortejo al alcanzar la madurez. En cualquier caso, varios investigadores han aplicado con éxito la reproducción artificial en individuos F1, demostrando el potencial de esta técnica, que aún debe ser optimizada para que su uso se pueda generalizar en el cultivo del lenguado [11, 12].

En la tabla 1 se recogen otros problemas que hoy en día persisten en este sector de la industria. Entre ellos, sobresale por un lado la aparición de malformaciones anatómicas y de fallos en la pigmentación de la piel durante la metamorfosis de las larvas, que podría estar relacionada con las condiciones nutricionales, la temperatura del agua y la duración del fotoperiodo. Asimismo, la incidencia de patologías en la fase del engorde es otra grave ocurrencia que posiblemente venga derivada de fallos en el control de las condiciones higiénicas y de temperatura y de la falta de estandarización de los piensos utilizados [1–3, 8].

Una elevada variabilidad de la talla de los ejemplares dificulta también el cultivo; para solucionarlo, se requiere un estudio extensivo de las bases genéticas y ambientales determinantes del tamaño con el fin de poder aplicar programas de selección genética que permitan conseguir individuos mejor adaptados a las necesidades de la cría intensiva. Por último, hay bastante dificultad a la hora del sexado de los peces por la falta de caracteres externos diferenciadores, un problema bastante grave dada la obligación de mantener una estricta relación macho/hembra de como mínimo 1:1 para asegurar el éxito de la fecundación. Este último punto pone en evidencia la actual falta de conocimiento sobre la influencia de factores externos tales como la temperatura en la diferenciación sexual del lenguado, así como de la urgente necesidad de identificar marcadores genéticos que faciliten la tarea del sexado [1–3, 8].

Tabla 1. Problemas que dificultan el cultivo a gran escala del lenguado senegalés junto a las causas sospechadas y posibles medidas a tomar en la industria, además de futuras líneas de investigación [1–3, 8].

Problemas en el cultivo	Posibles causas	Posibles soluciones
1. Dificultad de conseguir puestas regulares a partir de reproductores de primera generación en cautividad (F1)	- Ausencia de cortejo sexual en los machos F1	- Optimización de técnicas de reproducción artificial - Inducción por cambios en la temperatura del agua - Estudio de los factores responsables del fracaso reproductivo
2. Escaso volumen del eyaculado y alta variabilidad del mismo	- Propia biología de la especie	- Estudio de los factores determinantes de la calidad y cantidad del esperma
3. Malformaciones anatómicas y fallos en la pigmentación de la piel durante la metamorfosis	- Mal ajuste de factores externos como temperatura e iluminación	- Estudio de la influencia de estos parámetros sobre el desarrollo larvario
4. Aparición de enfermedades oportunistas durante el engorde	- Descuidos en las condiciones higiénicas y de temperatura de los tanques - Diferencias entre piensos de destete	- Mayor control sobre las condiciones de cultivo y uso de piensos de fórmula - Estudio de los requerimientos nutricionales - Desarrollo de vacunas
5. Elevada dispersión en el tamaño de los ejemplares	- Falta de información sobre el grado de influencia en la dispersión de talla de factores genéticos y factores sociales de interacción entre individuos	- Programas de mejora genética y de análisis poblacional del crecimiento
6. Mayor proporción de hembras que de machos	- Crecimiento más rápido en hembras - Dificultad de distinguir entre sexos fuera de la época de puesta - Maduración del macho poco evidente por el escaso volumen de eyaculado	- Estudio de la influencia de factores externos como la temperatura en la determinación sexual - Identificación de marcadores genéticos para la determinación sexual

4.3. Mecanismos de determinación sexual en vertebrados

Como se ha expuesto en el apartado anterior, dilucidar el mecanismo de determinación sexual en el lenguado senegalés conllevaría indudables ventajas. Por un lado, permitiría controlar la proporción macho/hembra en la población cultivada de una manera más eficiente, esencial para asegurar el éxito reproductivo. Este conocimiento también podría usarse para poner en marcha programas de selección genética con el objetivo de mejorar caracteres relacionados con el sexo, tales como el tamaño, de suma importancia [13]. Igualmente, ayudaría a la obtención de poblaciones de un único sexo, muy útiles cuando uno de los sexos crece más rápidamente que el otro y llega a alcanzar mayores tamaños, como es el caso en muchos peces incluyendo el lenguado, en el que al madurar las hembras más tarde son capaces de aprovechar esos nutrientes para el crecimiento [13].

En mamíferos y aves, el sexo viene estrictamente determinado en el momento de la fertilización por los cromosomas: en la mayoría de especies de mamíferos, los machos son el sexo heterogamético (sistema XX-XY) teniendo un cromosoma X y un cromosoma Y, que es parcial o completamente no homólogo al X, dado que en lugar de genes funcionales es rico en ADN repetitivo no codificante y transposones, por lo que es complicado que se dé recombinación entre los cromosomas sexuales. La diferenciación sexual primaria vendrá luego determinada por la activación del gen *sry* (*sex-determining region Y*) en el cromosoma Y, que desencadena una cascada de reacciones bioquímicas resultantes en la formación del fenotipo masculino. En cambio, en los vertebrados de sangre fría y entre ellos particularmente en los peces teleósteos, el proceso es mucho más complicado y variable, yendo de la unisexualidad (especies de solo hembras) al hermafroditismo pasando por el gonocorismo (que presenta el lenguado senegalés [14, 15]). En todos ellos, el fenotipo se encuentra regulado por varios mecanismos de determinación sexual que puede ser de dos tipos: genético y ambiental, que se pueden dar simultáneamente o no [14, 15].

Dentro de la determinación sexual genética, el sistema más estudiado son los cromosomas sexuales, que pueden ser distinguibles mediante técnicas citológicas (heteromórficos) o parecer idénticos (homomórficos). En ambos casos, un sexo siempre es heterogamético (con cromosomas sexuales diferentes) y el otro es homogamético (con dos copias idénticas de un mismo cromosoma sexual). Si el macho es heterogamético, el sistema recibe el nombre de XX-XY; mientras que si se trata de la hembra se denota sistema ZZ-ZW. Asimismo, los peces pueden tener sistemas más complejos que involucren a más de una pareja de cromosomas: en una única especie de pez, puede haber más de dos cromosomas sexuales al mismo tiempo e incluso coexistir dos tipos de cromosomas sexuales, lo que suele deberse a fusiones cromosómicas [15]. En vertebrados no mamíferos no se ha hallado un solo gen maestro regulador de la determinación sexual común como en mamíferos (con el gen *sry*), pero sí se han observado genes muy conservados que actúan de manera similar [16]. Estos genes se expresan de manera progresiva en las gónadas indiferenciadas y las llevan a desarrollarse hacia testículos u ovarios (gonadogénesis) [13].

Actualmente se cree que en la mayor parte de los vertebrados no mamíferos la diferenciación sexual femenina depende esencialmente de la estimulación de *cyp19a1* (gen de la enzima aromatasa, que convierte los andrógenos en estrógenos de manera irreversible) a través de un bucle de retroalimentación positiva en el que participa el factor de transcripción *foxl2*. Mientras tanto, la diferenciación sexual masculina parece darse por la inhibición de la expresión de *cyp19a1* lograda mediante la estimulación de *dmrt1*, un factor de transcripción altamente conservado en todos los taxones y que ha demostrado ser un inhibidor de *cyp19a1* en peces [17]. La regulación de la expresión de *dmrt1*, a su vez, puede darse por diferentes mecanismos tales como altas temperaturas (en el caso de peces donde la determinación sexual tenga un componente ambiental), un gen maestro regulador de la determinación sexual masculina o un gen asociado (como *sox9* en especies donde el fenómeno se atribuya completamente a la genética), o por los efectos aditivos de varios genes promotores del sexo masculino [17].

El primer gen maestro regulador de la determinación sexual descrito en peces se trata de *dmy* (también llamado *dmrt1bY*), presente en dos especies del género *Oryzias*, que surgió a partir de una duplicación del gen autosómico *dmrt1* previamente mencionado. Este último gen juega un papel vital en la determinación sexual, la diferenciación y el mantenimiento de órganos en una gran variedad de especies vertebradas, motivo por el que se le dedicará el próximo apartado [18].

El lenguado senegalés cuenta con un cariotipo de $2n=42$ cromosomas con un sistema XX/XY para la determinación sexual, sin haberse encontrado cromosomas heteromórficos o un gen regulador maestro por el momento [19]. No obstante, varios estudios citogenéticos sugieren que la primera de las tres parejas de cromosomas metacéntricos podría tratarse de una pareja de cromosomas proto-sexuales (que son cromosomas autosómicos portadores de un alelo determinante del sexo, que pueden exhibir o no todas las características de los cromosomas sexuales [20]) [21–25].

Esta hipótesis se basa en la evidencia de la acumulación de secuencias repetitivas en regiones donde existen posibles genes de determinación sexual (clúster *dmrt1-dmrt3-dmrt2*) y la disminución de la densidad genética del cromosoma metacéntrico, además del descubrimiento del origen de estos cromosomas, que habrían surgido por fusión Robertsoniana de dos parejas acrocéntricas, lo cual se trata de un suceso común en la evolución de los cromosomas sexuales [21–25]. Esto también explicaría las diferencias observadas con respecto a dos especies de pez plano estrechamente relacionadas con el lenguado senegalés (*Scophthalmus maximus* y *Scophthalmus rhombus*), que tienen un cariotipo de $2n=44$ cromosomas con dos parejas metacéntricas [26].

En cuanto a los factores ambientales que pueden influir sobre la determinación sexual se encuentran el pH del agua, la concentración de oxígeno, la tasa de crecimiento, la densidad de la población, la intensidad de la luz, el estatus social y más comúnmente, la temperatura del agua [13, 15]. En el caso de *S. senegalensis*, hay indicios acerca de la influencia de esta última sobre la diferenciación sexual [27].

4.4. Gen *dmrt1* en el lenguado senegalés y otras especies relacionadas

El gen de determinación sexual *dmrt1* (*doublesex and male abnormal-3 related transcription factor 1*) es un factor de transcripción que pertenece a una familia muy conservada en peces [28] caracterizada por un motivo común muy conservado de dominio de unión al ADN de dedos de zinc rico en cisteínas (llamado dominio DM). Como ya se ha comentado con anterioridad, se trata de un gen estrechamente relacionado con la regulación de la gonadogénesis masculina y la diferenciación de los testículos en casi todos los vertebrados, y especialmente, en peces gonocóricos [29].

El profundo interés por el papel de *dmrt1* en la determinación sexual de los peces nació con el descubrimiento ya mencionado de un homólogo suyo en el cromosoma Y de la medaka (*Oryzias latipes*), producto de una duplicación del gen autosómico *dmrt1*, al que se le llamó *dmy* [18, 30]. Este gen *dmy* resultó ser el regulador maestro dominante del desarrollo masculino, ejerciendo un papel similar a *sry* en mamíferos [18, 29, 30]. En este pez, las mutaciones que afectan a *dmy* provocan la inversión del sexo masculino al femenino; y la inducción del mismo mediante transgén es necesaria, y suficiente, para desencadenar el desarrollo masculino en hembras XX [30, 32, 33]. Estudios filogenéticos realizados a partir de secuencias proteicas concatenadas de genes relacionados con el sexo han hallado una gran proximidad entre la medaka y el lenguado senegalés, nuestra especie de interés [21].

Por otra parte, en la actualidad existen indicios de que *dmrt1* se podría tratar del regulador maestro en la lengua *Cynoglossus semilaevis*, otro pez plano muy cercano al lenguado que tiene un sistema cromosómico ZZ-ZW [33]. Curiosamente, el cromosoma Z de esta lengua, donde se ubica el gen *dmrt1*, presenta la misma ordenación de genes que un grupo de ligamiento concreto del lenguado (fenómeno denominado sintenia), lo que podría apoyar la hipótesis de *dmrt1* como regulador maestro en esta especie de lenguado, tal y como ocurre en la medaka y como parece ser el caso en *C. semilaevis* [34]. Igualmente, la secuencia proteica de *dmrt1* en *S. senegalensis* es la más cercana en términos de su evolución a las secuencias de esta especie de lengua, así como de *Scophthalmus maximus* y *Paralichthys olivaceus*, siendo todos pertenecientes al orden de los Pleuronectiformes [35]. Por todo ello, resultaría interesante estudiar con mayor profundidad el gen *dmrt1* en el lenguado senegalés por su posible rol en la regulación de la determinación sexual.

Un estudio reciente de Cross et al. [35] describió la estructura genómica y la secuencia nucleotídica del gen *dmrt1* a partir de ADNc (ADN complementario) de gónadas adultas (tanto masculinas como femeninas) en el lenguado senegalés. Se trata de un gen de unas 31.400 pb compuesto por 8 exones y 7 intrones (aunque como se comentará más tarde uno de los exones no es funcional, así que realmente habría 7 exones y 6 intrones) que contiene una duplicación intragénica de 10.000 pb que involucra parte de los exones I, II y III, como se aprecia en la figura 4, lo cual implica además que el dominio DM se encuentra parcialmente duplicado (en los exones I, II y II').

También se ha encontrado en la secuencia un elemento SINE (*short interspersed nuclear element*, o elemento nuclear intercalado corto) flanqueando la segunda región duplicada, que ha sido propuesto como el factor responsable de esta duplicación mediante barajado de exones.

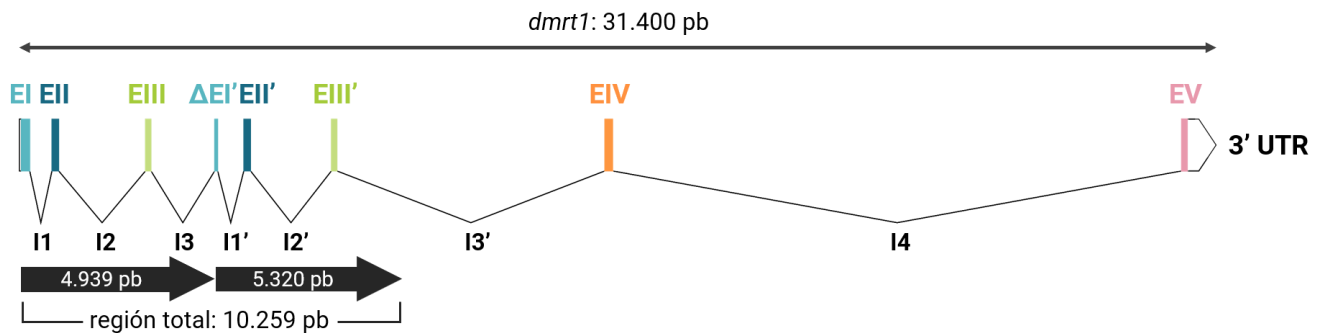


Figura 4. Esquema a escala del gen *dmrt1* en el genoma de *S. senegalensis*, donde EI-EV (en colores) denota los exones y I1-I4 (en negro) representa los intrones. Los extremos en negro son las regiones 5'UTR (de longitud indeterminada) y 3'UTR. Fuente: modificada a partir de Cross et al. [35].

Tras alinear los exones duplicados entre sí, se vio que el exón $\Delta I'$ alineaba perfectamente con la segunda parte del I (aunque sólo cubría un 35% de la secuencia), pero carecía de un ORF (*open-reading frame*, o marco de lectura abierto) correcto, por lo que $\Delta I'$ no se trataría de un exón funcional. Por su parte, los exones II y II' mostraron una secuencia casi idéntica, mientras que el alineamiento entre los exones III y III' resultó en 8 SNPs (*single-nucleotide polymorphisms*, o polimorfismos de un solo nucleótido). Por tanto, la duplicación comienza exactamente en el exón $\Delta I'$ y se extiende hasta parte del intrón 3', tal y como se esquematiza en la figura 4. Al analizar las secuencias de la región comprendida entre los exones I y IV, se observaron numerosos polimorfismos producidos por procesamiento alternativo de los exones [35].

El procesamiento del ARNm precursor (o pre-ARN) es una etapa reguladora crucial en la vía de expresión de genes, puesto que es donde se eliminan los intrones y se ligan los exones para formar el ARNm (ARN mensajero) maduro. Al incluir diferentes exones en el ARNm, el procesamiento alternativo da lugar a la generación de distintas isoformas a partir de un único gen; este es uno de los motivos por los que en los organismos eucariotas hay mucho menos genes que proteínas diferentes. Este rol generador de diversidad podría verse aumentado con la aparición de exones alternativos, que entre otras formas puede darse por barajado de exones, como parece ser en el caso del gen *dmrt1*. En este proceso, se inserta un exón nuevo en un gen ya existente o se duplica un exón del mismo gen [36]. Se estima que entre el 8-17% de los genes de un genoma eucariota típico tienen alguna región exónica y/o intrónica duplicada [37].

Dmrt1 es transcrito a un ARNm de 2.027 pb de longitud y 1.206 pb de ORF, que codifica para una proteína de 402 aminoácidos. Se han descrito 6 transcritos de ARNm diferentes, en los que el exón I siempre está presente y el exón IV aparece en todos los clones menos uno. Aunque demuestran una gran variabilidad en el procesamiento alternativo de la expresión del gen, no pudieron ser relacionadas con el sexo de los individuos. Aun así, se piensa que podrían constituir diferentes dianas para factores participantes en la regulación sexual [35].

5. OBJETIVOS

Dado el interés económico del lenguado y la dificultad que supone su reproducción en cautividad, el objetivo principal propuesto en el presente trabajo se centra en profundizar en la estructura del gen *dmrt1* en el lenguado senegalés (*Solea senegalensis*), poniendo a prueba su posible relación con el sexo en términos de variabilidad y expresión génica. Esto ayudaría a dilucidar si realmente se trata de un gen implicado en la determinación sexual en esta especie, puesto que se ha observado una duplicación en el mismo que podría sugerir un papel similar al que presenta en otros teleósteos. Para lograrlo, se han planteado diferentes objetivos específicos:

1. Confirmar la estructura genómica de *dmrt1*, descrita recientemente, y analizar su secuencia mediante amplificación por PCR larga de todas las regiones intrónicas de múltiples individuos y posterior secuenciación.
2. Digerir ciertas regiones intrónicas con enzimas de restricción para buscar polimorfismos que pudieran ser sexo-específicos (concretamente el intrón 3, por estar flanqueado por exones donde la duplicación no es exacta, lo que permite el diseño de *primers* específicos).
3. Determinar la eficiencia de *primers* específicos para PCR cuantitativa (qPCR) que permitan la cuantificación del ARNm de *dmrt1* a través de regiones conservadas (exones I y IV, por ser comunes a prácticamente todos los transcritos de ARNm detectados hasta la fecha) y cuantificar la expresión génica de *dmrt1* en individuos de distinto sexo, madurez y origen por qPCR.

6. MATERIALES Y MÉTODOS

A modo de resumen, la figura 5 muestra los bloques de trabajo en los que se han agrupado las tareas de laboratorio que se van a ir desarrollando a lo largo de este epígrafe, entre otras: (1) preparación de muestras de ADN (genómico y complementario); (2) amplificación con 7 parejas de *primers*; (3) digestión del intrón 3 de *dmrt1* con enzimas de restricción para análisis de polimorfismos; y (4) puesta a punto de dos parejas de *primers* para cuantificar el ARNm (exones I y IV) de *dmrt1* mediante qPCR.

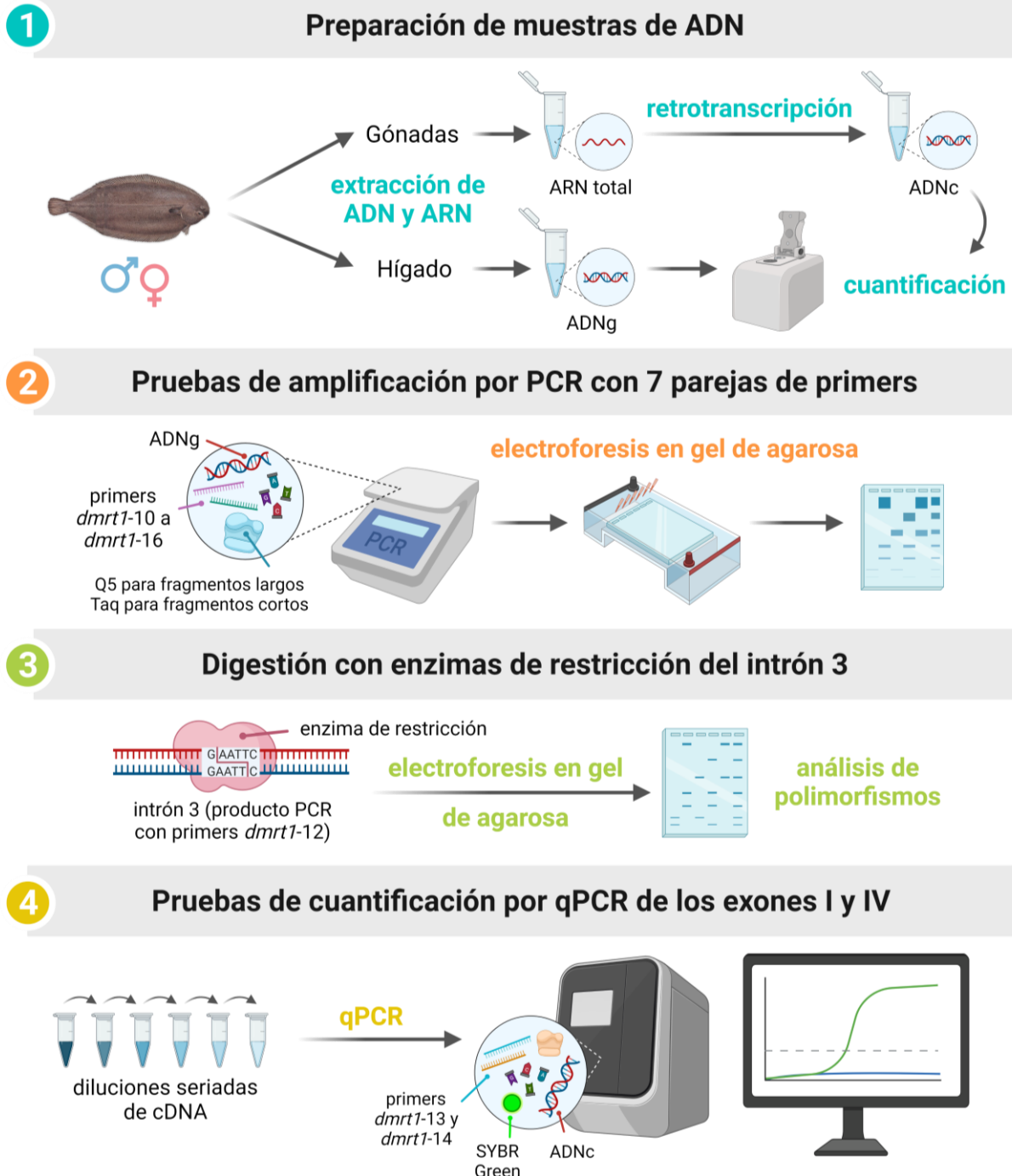


Figura 5. Resumen del procedimiento llevado a cabo en el laboratorio para la realización de este trabajo. Creada con BioRender.com.

6.1. Diseño de *primers*

Tras obtener la secuencia completa del gen *dmrt1*, resultante de unir los exones e intrones, se procedió al diseño de *primers* (también llamados cebadores, oligonucleótidos o simplemente oligos) empleando el software online Primer3 (v. 0.4.0) [38, 39] disponible en <https://bioinfo.ut.ee/primer3-0.4.0/>.

Figura 6. Interfaz del software online Primer3 (v. 0.4.0).

Para su diseño con Primer3 se tuvo en cuenta una serie de parámetros, siendo los siguientes los más destacables:

- **Regiones excluidas:** en todos los casos, se trató de excluir las regiones repetidas provenientes de la duplicación parcial del gen, puesto que podrían dar lugar a amplificación múltiple. Además, se procuró que las secuencias de los *primers* pertenecieran a los exones correspondientes a la región a estudiar, en lugar de a los intrones, dado que los primeros suelen estar mucho más conservados evolutivamente debido a la presión selectiva por codificar para proteínas. En ese mismo sentido, el hecho de que el código genético sea degenerado y redundante nos llevó a diseñar *primers* que terminaran en su extremo 3' en primera o segunda base del codón concreto. Se tomó esta medida de precaución porque la existencia de polimorfismos (que son más probables en la tercera posición del codón, de nuevo, a causa de que los codones que codifican para un mismo aminoácido suelen compartir las dos primeras bases) podría disminuir la especificidad de la pareja de *primers* resultante. Las regiones excluidas por no cumplir estos requisitos fueron delimitadas en el programa con los símbolos "<>".
- **Tamaño del producto:** para las parejas destinadas a amplificación larga del gen, así como del intrón 3, se buscó maximizar la longitud del amplicón, de forma que tuviera la mayor cobertura posible. En contraposición, para los *primers* a utilizar en qPCR se buscó que el producto amplificado tuviera un tamaño entre 100-200 pb, porque por debajo de este límite es difícil distinguirlo de dímeros de *primers* y por encima disminuye la eficiencia de la técnica con SYBR Green [40].

- **Longitud del *primer*:** típicamente, la longitud óptima de un *primer* está entre 18-24 pb, así que se intentó que los *primers* diseñados tuvieran alrededor de 20 nucleótidos puesto que la longitud influye de forma importante sobre la Tm [41].
- **Temperatura de fusión (Tm) del *primer*:** la Tm (*melting temperature*, o temperatura de fusión), que es la temperatura a la que la mitad de las moléculas del *primer* está unido a su complemento, debe estar entre 57-63°C. Además, la diferencia entre la Tm teórica del *primer forward* (Fw) y del *reverse* (Rev) no puede sobrepasar los 5°C, por lo que se impuso en el programa una diferencia máxima de 1-2°C que en ocasiones tuvo que ser incrementada. Concretamente, para los *primers* a usar en qPCR con SYBR Green se recomienda minimizar la longitud lo máximo posible, con una Tm óptima de 63°C [42].
- **Contenido en GC:** para garantizar una unión estable entre el *primer* y el molde, los *primers* deben contener entre un 40-60% de guanina y citosina (GC) [41].
- **Extremo 3' GC:** la presencia de 1 o 2 bases G o C en las últimas cinco bases del extremo 3' de los *primers* ayuda a promover la unión específica del *primer* al molde, ya que estos nucleótidos se unen entre sí a través de tres enlaces por puente de hidrógeno (en lugar de dos) y forman uniones fuertes y estables. No obstante, se debe evitar que haya 3 o más bases GC en este extremo, puesto que podría aumentar excesivamente la Tm y llevar a la formación de dímeros de *primers* [41].
- **Secuencias poli-X:** en este trabajo se limitó el máximo número de nucleótidos repetidos seguidos a 3 (poli-X: AAA, TTT, GGG o CCC) porque una cantidad superior puede propiciar una amplificación inespecífica (*mispriming*) [40].
- **Especificidad:** siempre es necesario verificar que los *primers* diseñados no hibriden con otras partes del genoma en estudio que no sean el gen de interés [40]. Para ello, se empleó Primer-BLAST utilizando las bases de datos disponibles, pero al no haberse secuenciado por completo el genoma de *S. senegalensis* (cuando se hizo el diseño de *primers*), resultó prácticamente imposible evaluar con certeza su especificidad. Por eso, no se pudo realizar una PCR *in silico* de los mismos hasta después de haber finalizado los experimentos de laboratorio, cuando estuvo disponible la secuencia del genoma del lenguado, como se describirá en el último apartado.
- **Autocomplementariedad:** hay que evitar *primers* que formen dímeros consigo mismos (homodímeros) y horquillas (*hairpins*), especialmente para los *primers* de qPCR, puesto que la molécula de SYBR Green se une a cualquier doble hélice de ADN indiscriminadamente. Tratando de minimizar el valor de la autocomplementariedad lo máximo posible, se fijó inicialmente en 2 y se fue incrementando o disminuyendo según los *primers* resultantes [40].
- **Complementariedad 3':** también hay que evitar que se formen dímeros de *primers* entre el Fw y el Rev (heterodímeros). Como norma general, hay que descartar cualquier pareja que sea homóloga para 3 o más pb en sus extremos 3'. Igual que para el parámetro anterior, se comenzó usando 1 como valor y se fue variando a fin de minimizar la complementariedad 3' [40].

Una vez diseñadas y elegidas las parejas, se encargaron a la empresa MacroGen. Se añadió 250 μL de agua estéril a cada uno de los tubos con los *primers* liofilizados para crear soluciones madre (o stock) con concentración final de 100 μM . A continuación, se realizaron diluciones 1:10 para conseguir stocks de trabajo de 10 μM . Esto se hizo para ayudar a evitar la contaminación de la solución madre, ya que reduce el número de ciclos de congelación y descongelación de la misma.

Hay que destacar que a la hora de diseñar los *primers* para este trabajo, se eligieron las parejas exclusivamente en base a los resultados de Primer3; pero en vista de la gran cantidad de amplificación inespecífica y la falta de productos esperados en las PCR, se optó por hacer una valoración *a posteriori* de los dos últimos parámetros relacionados con la complementariedad.

Para tal fin, se realizó un análisis de las estructuras secundarias que pueden formar tanto los *primers* como los amplicones. Mediante el software online OligoAnalyzer 3.1 de Integrated DNA Technologies (<https://eu.idtdna.com/calc/analyzer>) se obtuvieron los valores de estabilidad termodinámica, dados a través de la energía libre de Gibbs (ΔG), para las posibles estructuras que podrían formarse como resultado de la interacción entre los *primers* tales como homodímeros, heterodímeros y horquillas [39].

Los parámetros se ajustaron en la mayor medida de lo posible a las concentraciones reales que se dieron en las reacciones (concentración de *primers*, Na^+ , Mg^{++} y dNTPs) y a las temperaturas de hibridación (este último para la formación de horquillas).

Un *primer* es más apto cuanto mayor y más positiva sea la ΔG de las posibles estructuras no deseadas, ya que así no serían termodinámicamente estables. Si bien no existe un límite concreto, en la herramienta se recomienda que tengan valores mayores de -9 kcal/mol (aunque, de nuevo este número es aproximado, porque no existe un valor real que garantice la eficacia de la PCR; los *primers* deben ser comprobados en el laboratorio). En el caso de que se formen horquillas, debe evitarse que se den en el extremo 3' del *primer* y que tengan muchos puntos de unión [40].

Por otro lado, para el análisis de la formación de estructuras secundarias a partir de los amplicones se empleó el servidor web de mFold, disponible a través de <http://unafold.org/mfold/applications/dna-folding-form.php> (solamente para los *primers* de qPCR, puesto que el software tiene un límite de secuencia de 2.400 pb) [43]. Con él se verificó que las estructuras secundarias de los amplicones tuvieran una T_m menor que la temperatura de hibridación (también llamada de anillamiento), que sería de 60°C, pues si no fuera así, estas posibles estructuras podrían no haberse desnaturalizado antes del paso de hibridación de los *primers* [40]. Como temperatura de plegamiento (*folding temperature*) se usó la de hibridación (60°C) y se trató de ajustar las condiciones iónicas a las reales (concentración de Na^+ y Mg^{++}), dejando el resto de parámetros por defecto.

6.2. PCR *in silico*

Para comprobar la especificidad de las 7 parejas de *primers* diseñados, se realizó una PCR *in silico* con el ensamblaje del genoma publicado después de haber realizado los experimentos [44], usando el paquete exonerate de iPCress (v. 2.4.0-4) [45].

Se utilizó el comando `ipccress -i primers.txt -s sse.fasta --mismatch 4 > resultados.txt`, donde “primers.txt” es un archivo de texto en el que se especifica la ID del experimento, las secuencias de los *primers* Fw y Rev de una misma pareja y los tamaños de producto mínimo y máximo; y donde “mismatch 4” indica que el número máximo de emparejamientos incorrectos permitidos es hasta 4. En ningún caso se definió tamaño mínimo; para los *primers* de amplificación larga (*dmrt1*-10, 11, 12, 15 y 16, que como se verá en la tabla 12 de resultados dan productos de alrededor de 15.000 pb, por lo que en la amplificación se usaría una polimerasa Q5 capaz de dar fragmentos más largos) se definió un tamaño máximo de 20.000 pb, mientras que para *primers* de amplificación corta (*dmrt1*-13 y 14, que al amplificar menos de 200 pb se usarían con una polimerasa Taq estándar) se usó 3.000 pb.

6.3. Extracción de ADN genómico

El ADN genómico fue extraído a partir de muestras de hígado de tres machos y tres hembras. En la extracción se siguió el protocolo esquematizado en la figura 7, comenzando por cortar las muestras en pequeñas piezas empleando un bisturí, utilizando entre 0,1-0,2 gramos de tejido.

A continuación, se introdujeron en tubos Eppendorf de 2 mL y se añadió 1 mL de buffer de lisis CTAB con un 2% de Antifoam B. Se añadieron dos bolas cerámicas a cada uno de los tubos y se vortearon hasta que el tejido estuviera completamente disgregado. Para degradar las proteínas de membrana y otras proteínas que se encuentren unidas al ADN, se incubaron las muestras con 2 µL de proteinasa K (con una concentración final en la solución de 100 µg/mL) en el termobloque Tembloc (J.P. Selecta) durante aproximadamente 1 hora.

Luego se añadió un volumen (600 µL) de cloroformo:alcohol isoamílico (24:1) y se mezcló por completo en el vórtex, para después centrifugar los tubos 7 minutos a 13.000 rpm. Habiendo separado la fase acuosa superior (que contiene el ADN) de la fase orgánica inferior (con lípidos, proteínas y otras impurezas), se transfirieron 500 µL de la fase acuosa a tubos nuevos, tratando de no tocar la interfase. Para precipitar la fase acuosa se emplearon 250 µL de NaCl 5M y 500 µL de isopropanol, que tras mezclarse bien se centrifugó durante 10 minutos a 13.000 rpm. Se descartó el sobrenadante y se lavó el *pellet* (precipitado) con 500 µL de etanol al 70%. Por último, el *pellet* ya seco se resuspendió en 100 µL de agua estéril.

Para verificar la integridad de las seis muestras, se realizó una electroforesis en gel de agarosa al 1% en TBE 1X, dejándolo correr durante 60 minutos a 77V. En cada pocillo se cargó 1 µL de las muestras de ADN_g, mezclado previamente con 1 µL de gTPbio Bluxyo Safe (usado como alternativa segura para teñir el ADN, en lugar de bromuro de etidio). Como marcador de pesos moleculares se utilizó Lambda DNA/HindIII. Tras finalizar la electroforesis, se visualizó el ADN con una cámara transiluminadora de luz ultravioleta (Syngene Ingenius 3) y con el software GeneSys.

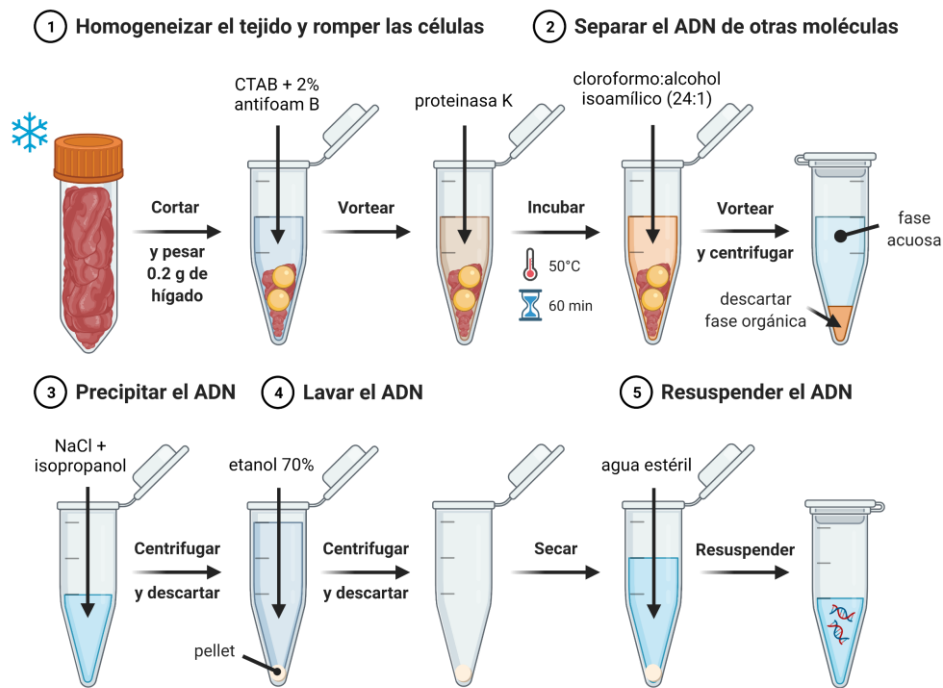


Figura 7. Protocolo de extracción de ADN genómico a partir de muestras de hígado. Creada con BioRender.com.

Posteriormente, la cuantificación y determinación de la pureza del ADN_g extraído se llevó a cabo por medio de espectrofotometría, midiendo la absorbancia a 230, 260 y 280 nm (máximos de absorbancia para ADN, ARN y proteínas, respectivamente) con el NanoDrop 2000c (Thermo Scientific). El cociente entre las absorbancias a 260 y 280 nm (A_{260}/A_{280}) para ADN puro es de ~ 1.8 , mientras que para ARN puro es de ~ 2 (siendo valores que se desvían de este rango resultado de contaminación con proteínas). Por otra parte, el cociente A_{260}/A_{230} para ácidos nucleicos puros debería ser de 2-2.2 (pudiendo alejarse de estos valores por presencia de carbohidratos u otros compuestos empleados en el protocolo de extracción) [46]. Una vez cuantificadas las muestras de ADN_g, se hicieron las diluciones apropiadas para tener stocks de trabajo de 100 ng/ μ L y las soluciones madre se almacenaron a -20°C.

6.4. Amplificación por PCR

Todas las amplificaciones por PCR estándar realizadas a lo largo de este trabajo (tanto de fragmentos cortos como largos) emplearon uno de los siguientes tres termocicladores: SimpliAmp Thermal Cycler (Applied Biosystems), T100 Thermal Cycler (Bio-Rad) o GeneAmp PCR System 2700 (Applied Biosystems), según la disponibilidad del laboratorio. Para los gradientes de temperatura se utilizó el segundo equipo.

Para la amplificación larga, se diseñaron distintas parejas de *primers* alternativas para amplificar el gen *dmrt1* por completo, dividiéndolo en dos mitades de unas 15.000 pb cada una: del exón I al IV (*primers dmrt1-10 y 15*) y del exón V al VI (*primers dmrt1-11 y 16*). Debido a la larga longitud de los fragmentos a amplificar, se utilizó la ADN polimerasa Q5 High-Fidelity (New England Biolabs), específica para este tipo de PCR. Las reacciones se prepararon de acuerdo a los volúmenes recomendados por el fabricante, que son los enumerados en la tabla 2, preparando master-mix en función del número de reacciones.

Tabla 2. Volúmenes empleados para la PCR estándar de fragmentos largos con la polimerasa Q5.

Componente	Vol. (1 rxn) (μL)	Concentración final
Buffer de reacción Q5 (5X)	5	1X
dNTPs (10 mM)	0,5	200 μM
Primer Fw (10 μM)	1,25	0,5 μM
Primer Rev (10 μM)	1,25	0,5 μM
ADN molde (100 ng/μL)	2	200 ng
Q5 High-Fidelity	0,25	0,02 U/μL
Agua libre de nucleasas	14,75	
Total	25	

Casi todas las PCR de fragmentos largos se prepararon con estos volúmenes, excepto las realizadas antes de la extracción de ADN genómico en las que se empleaba 0,5 μL de ADN, pues la muestra era de 200 ng/μL, usando por tanto 16,25 μL de agua; y aquellas en las que se usaron 5 μL de Enhancer Q5 High GC 5X por reacción para optimizar la amplificación (para fragmentos de alto %GC), añadiendo por tanto un menor volumen de agua.

En cuanto a las condiciones de tiempo y temperatura en el termociclador también se usaron las sugeridas por el proveedor, en la tabla 3. La temperatura de hibridación fue calculada con la calculadora de T_m de la empresa (<http://tmcalculator.neb.com>). El tiempo de extensión para la PCR-10 (con los *primers dmrt1-10*) se varió con fin de optimizar la reacción, de 7,5 a 11 y a 12,5 minutos.

Tabla 3. Condiciones de tiempo y temperatura para la PCR estándar de fragmentos largos con la polimerasa Q5.
*La temperatura de hibridación para cada pareja de *primers* fue determinada con la calculadora de T_m de NEB.

Paso	Temperatura	Tiempo
Calentamiento previo	98°C	3 minutos (saltar)
Desnaturalización previa	98°C	30 segundos
Desnaturalización	98°C	10 segundos
Hibridación/anillamiento	*50-72°C	30 segundos
Extensión		50 segundos/Kb
Extensión final	72°C	2 minutos
Hold	12°C	Infinito

Por otro lado, dado el tamaño de producto PCR esperado con los *primers dmrt1-12* (de 2.522 pb), se probó a amplificar tanto con la polimerasa Q5 (siguiendo los volúmenes y condiciones de termociclador de las tablas 2 y 3) como con la ADN polimerasa NZYtaq II (NZYTech), que debería funcionar hasta las 6.000 pb. Con el protocolo Q5 se probó asimismo la amplificación con Enhancer Q5 High GC 5X, variando por tanto el volumen de agua. También se realizó un gradiente de temperaturas para el paso de hibridación, de 60 a 72°C, y una vez hallada la temperatura óptima (de 71.1°C) se empleó esa misma para el resto de experimentos.

Para la amplificación con NZYtaq II se usaron los volúmenes y condiciones indicados en las tablas 4 y 5. Como temperatura de hibridación se empleó la menor T_m (dada por MacroGen) de la pareja de *primers* menos 3°C [47]. Después de probar con las condiciones de termociclador recomendadas para la polimerasa, se probaron a utilizar las de Q5 y desde entonces se usaron esas por dar mejores resultados, utilizando todavía NZYtaq II. Finalmente, tras comparar las amplificaciones con Q5 y NZYtaq II nos decantamos por utilizar NZYtaq II (con las condiciones de tiempo y temperatura propias de Q5) para preparar las muestras de los 6 individuos a digerir con enzimas de restricción en un experimento posterior.

Componente	Vol. (1 rxn) (μL)	Concentración final
Buffer de reacción (10X)	5	2X
MgCl ₂ (50 mM)	1,5	3 mM
dNTPs (10 mM)	0,5	200 μM
Primer Fw (10 μM)	1	0,4 μM
Primer Rev (10 μM)	1	0,4 μM
ADN molde (100 ng/μL)	1,5	150 ng
NZYTaq II (5 U/μL)	0,4	0,08 U/μL
Agua libre de nucleasas	16,6	
Total	25	

Tabla 4. Volúmenes empleados para la PCR estándar de fragmentos cortos con la polimerasa NZYTaq II.

Tabla 5. Condiciones de tiempo y temperatura para la PCR estándar de fragmentos cortos con la polimerasa NZYTaq II. *Como temperatura de hibridación para cada pareja de primers se utilizó la menor T_m de la pareja de primers menos 3°C.

Paso	Temperatura	Tiempo
Calentamiento previo	95°C	3 minutos (saltar)
Desnaturalización previa	95°C	3 minutos
Desnaturalización	94°C	30 segundos
Hibridación/anillamiento	*T _m -3°C	30 segundos
Extensión	72°C	30 segundos/Kb
Extensión final	72°C	7 minutos
Hold	12°C	Infinito

Como paso previo a la qPCR se hicieron pruebas preliminares para los primers *dmrt1-13* y *14* (que amplifican los exones I y IV, respectivamente), tanto con ADN_g como con ADN_c en ambos casos (aunque en qPCR sólo se emplearía ADN_c).

En ambos casos se utilizó la polimerasa NZYTaq II con los volúmenes y condiciones mencionados anteriormente en las tablas 4 y 5, ya que los productos PCR son cortos (inferiores a 200 pb). Para los primers *dmrt1-14* se realizó un gradiente de temperaturas de hibridación de 53 a 65°C con uno de los individuos, encontrando una temperatura óptima de 64.1°C que se usaría en adelante.

6.5. Electroforesis en gel de agarosa

La electroforesis en gel de agarosa fue empleada tanto para la verificación de productos de PCR y qPCR como para el análisis tras la digestión con enzimas de restricción del producto PCR, así como para verificar la integridad del ADN genómico extraído.

Para preparar los geles, se pesaba la agarosa en polvo y se disolvía en el volumen adecuado de tampón TBE 1X (Tris-borato-EDTA) calentándolo, a concentraciones finales de 1-1,5% (dependiendo del tamaño de las bandas a resolver: 1,5% para productos de poco tamaño y 1% para amplicones mayores, como los esperados al usar las parejas de primers *dmrt1-10*, *11*, *15* y *16*).

Las muestras a ser cargadas se mezclaban previamente con distintos volúmenes de gTPbio Bluxy Safe (que contiene azul de bromofenol, Blu; cianol de xileno FF, Xy; y naranja G, O), una alternativa más segura al bromuro de etidio (EtBr), que podría ser mutagénico. Se usaron cubetas de tamaño pequeño y mediano con tampón TBE 1X y se emplearon condiciones de voltaje constante, entre 60-120V (según las necesidades del experimento).

Uno de los geles se tiñó en bromuro de etidio (sumergiéndolo en una cubeta) durante 20 minutos tras haber corrido, para asegurar una correcta visualización de las bandas, y otros geles se tuvieron que preparar con 0,5 µg/ml de bromuro de etidio, utilizando tampón de carga en lugar de Bluxyo (para dar densidad a la muestra y que hacer que se sumerja en el fondo del pocillo).

Los geles eran introducidos en la cámara transiluminadora de luz ultravioleta Syngene Ingenius 3, empleándose el programa GeneSys para capturar las imágenes. Para estimar los tamaños de las bandas, estas se comparaban con marcadores de pesos moleculares como Lambda DNA/HindIII o HyperLadder II (también llamado HyperLadder 50bp), según si el tamaño de producto PCR esperado fuera mayor o menor.

6.6. Purificación de producto PCR-10 y escisión y reamplificación de bandas

En las pruebas de PCR con la pareja de *primers dmrt1-10* se observó una gran cantidad de bandas en los geles en lugar del único producto esperado, que debería ser de 15.470 pb. A fin de averiguar el motivo de esta amplificación inespecífica y en qué región del genoma podría estar sucediendo, se intentó cortar una de las bandas desconocidas (alrededor de 1.800 pb) para ser secuenciada (de dos individuos distintos: macho 1 y hembra 2).

Antes de cargar el producto PCR en el gel y cortar la banda se hizo una purificación del mismo para eliminar el exceso de reactivos de PCR (*primers*, buffers, dNTPs y enzimas), permitiendo que la escisión fuera lo más limpia posible. El protocolo seguido (ilustrado en la figura 8) fue el de purificación de producto PCR (*PCR clean-up*) del kit comercial NucleoSpin Gel and PCR Clean-up (Macherey-Nagel). Se comenzó mezclando 125 µL de producto PCR de cada uno de los dos individuos (1 volumen) con 250 µL de buffer NTI (2 volúmenes), cuyo contenido en sales caotrópicas crea un entorno hidrofóbico alrededor del ADN que permite su unión a una columna de sílice y su posterior separación.

Después, se cargó 375 µL de la muestra recién mezclada en un nuevo tubo a través de la columna de sílice provista por el kit. Tras centrifugar los tubos con sus respectivas columnas a 11.000 g durante 30 segundos, el ADN quedaría unido a la columna de sílice y se descartó el sobrenadante, volviendo a colocar la columna sobre el tubo de recogida. A continuación, se eliminó la posible contaminación lavando la columna con 700 µL del buffer NT3 (con contenido en etanol) y centrifugando, de nuevo, a 11.000 g durante 30 segundos. Una vez descartado el sobrenadante y colocada la columna sobre el tubo, se centrifugó otra vez a 11.000 g durante 1 minuto para eliminar el exceso de buffer NT3. La elución del ADN de la columna se hizo añadiendo 20 µL de buffer NE (una sustancia ligeramente alcalina y con un bajo contenido en sales) e incubando las muestras a temperatura ambiente durante 1 minuto. Como último paso, se centrifugó a 11.000 g durante 1 minuto y se desechó la columna, quedando el producto PCR purificado disuelto en el tubo.

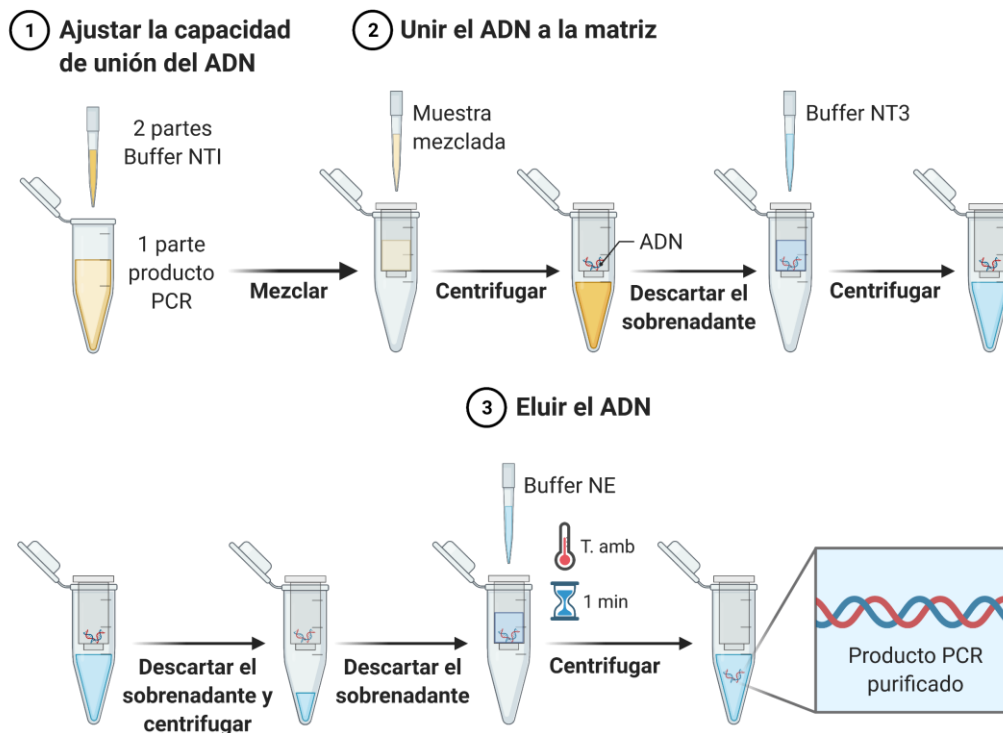


Figura 9. Protocolo de purificación de producto PCR del kit NucleoSpin Gel and PCR Clean-up (Macherey-Nagel). Creada con BioRender.com.

Para cortar la banda de 1.800 pb a ser secuenciada, se siguió el protocolo de extracción de ADN a partir de geles de agarosa (*DNA extraction from agarose gels*) del kit NucleoSpin Gel and PCR Clean-up (Macherey-Nagel) representado en la figura 9. El procedimiento es similar al anterior, con la diferencia de que antes hay que calentar el fragmento de agarosa para que se disuelva y poder extraer así el ADN. El gel, en el que previamente habían corrido las muestras purificadas de producto PCR, se introdujo en un transiluminador UV de sobremesa (Bioblock Scientific) y con ayuda de un bisturí estéril se cortaron dos fragmentos correspondientes a la banda de interés (uno para cada individuo, macho y hembra) que posteriormente se introdujeron en tubos Eppendorf. Durante la escisión, se trató de minimizar la exposición del gel a la radiación UV para evitar daños a la muestra.

Habiendo pesado los dos fragmentos de gel, se añadió 200 μ L de buffer NTI en el tubo por cada 100 mg de gel y se incubaron las muestras durante 10 minutos a 50°C en el termobloque, vorteoando cada 3 minutos hasta que los fragmentos se disolvieran en su totalidad. A partir de este paso, el procedimiento seguido fue idéntico al anterior.

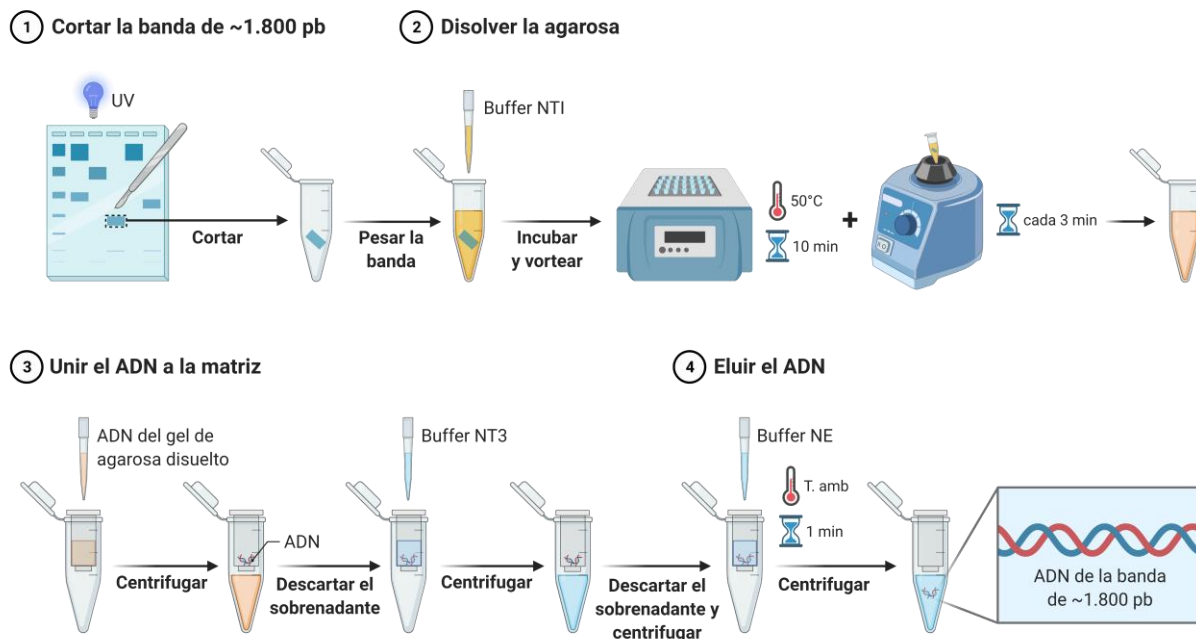


Figura 9. Protocolo de extracción de ADN a partir de geles de agarosa del kit NucleoSpin Gel and PCR Clean-up (Macherey-Nagel). Creada con BioRender.com.

Después de probar con NZYtaq II y Q5 y distintos tiempos de extensión, la reamplificación de la banda se llevó a cabo con la polimerasa Q5, siguiendo las condiciones de volúmenes indicadas en la tabla 2 (con la diferencia de que se usaron 1 y 5 μ L de ADN de la banda para cada individuo y por reacción) y las de termociclador de la tabla 3 (usando 12,5 minutos de extensión, que fue el tiempo utilizado en la PCR de cuyo gel se cortó la banda de 1.800 pb).

6.7. Digestión con enzimas de restricción del intrón 3

Con el objetivo de determinar las enzimas de restricción más aptas a ser utilizadas durante la digestión, se creó un mapa de restricción para el amplicón resultante de la amplificación con los *primers dmrt1-12* (correspondiente al intrón 3) y con un tamaño de 2.522 pb). Este fue realizado con el programa BioEdit (v. 7.2.5) utilizando la herramienta *Restriction Map*.

Dejando el resto de parámetros por defecto, se optó por mostrar las enzimas de restricción con dianas de 4, 5 y 6, incluyendo isoesquizómeros (enzimas con la misma secuencia de reconocimiento). Una vez generada la lista de enzimas de restricción con diana en la secuencia de interés, se seleccionaron en función del número y tamaño de fragmentos generados (idealmente, cortando 2 o 3 veces daría lugar a 3 o 4 fragmentos, respectivamente, que deberían tener tamaños adecuados para su visualización en gel y poder ser diferenciados entre sí) y de su disponibilidad comercial.

Para la digestión de los productos PCR, se siguieron las instrucciones de los proveedores de las enzimas de restricción *DraI* (New England Biolabs) y *PstI* (Thermo Scientific). Al tratarse de digestiones simples (usando una sola enzima por reacción), se trabajaba con 6 muestras de producto PCR de cada uno de los individuos en cada caso. Los volúmenes utilizados se recogen en la tabla 6.

Tabla 6. Volúmenes necesarios de reactivos para la digestión de producto PCR con las enzimas *DraI* y *PstI*, para cada reacción.

Volúmenes (μL)	<i>DraI</i>	<i>PstI</i>
Producto PCR (μL)	18	18
Buffer (μL)	2 (CutSmart)	3,5 (0)
Enzima (μL)	0,6	2,5

Las mezclas resultantes se incubaron durante 2,5 horas en el termociclador T100 Thermal Cycler (Bio-Rad) a la temperatura óptima de actividad de las enzimas, que es de 37°C en ambos casos. Cada muestra se mezcló con 4 μL de Bluxyo y fueron cargadas en un gel de agarosa al 1,5%, dejándolo correr a 80V durante 120 minutos.

6.8. Retrotranscripción de ARN a ADNc

Nos fueron facilitadas 6 muestras de ARN total que habían sido extraídas y purificadas a partir de tejido gonadal mediante el kit comercial miRNeasy Micro Kit (Qiagen) usando el protocolo de purificación de ARN total a partir de tejido animal (*Purification of total RNA, including miRNA, from animal tissues*).

La retrotranscripción del ARN a ADNc se hizo siguiendo las instrucciones del kit comercial QuantiTect Reverse Transcription (Qiagen), manteniendo las muestras en hielo en todo momento y utilizando puntas de pipeta con filtro para minimizar el riesgo de contaminación, como se muestra en la figura 10. El primer paso fue la eliminación del ADN genómico, usando los volúmenes recogidos en la tabla 7 de buffer gDNA Wipeout, ARN y agua estéril libre de ARNasas hasta un volumen de 14 μL.

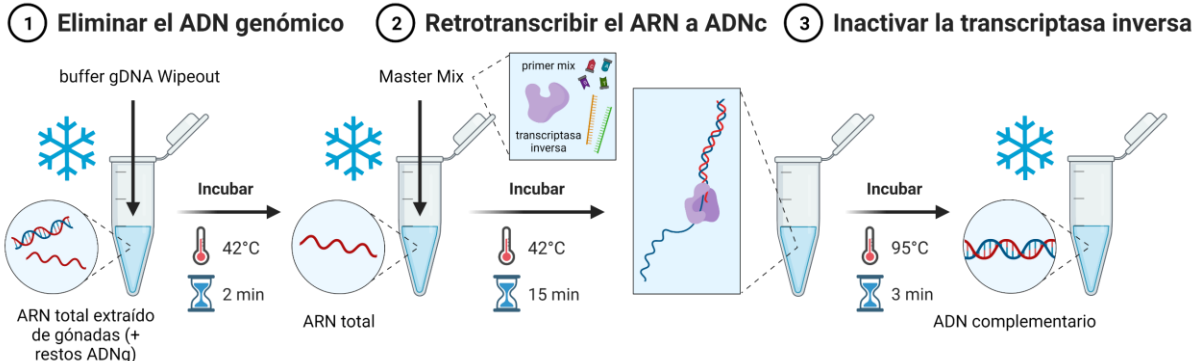


Figura 10. Protocolo de retrotranscripción de ARN a ADNc del kit QuantiTect Reverse Transcription (Qiagen). Creada con BioRender.com.

Ya que las puntas de pipeta con filtro de las que se disponía limitaban el volumen pipeteado a un mínimo de 2 μL, previamente se tuvieron que realizar diluciones apropiadas para no exceder 1 μg.

Tabla 7. Volúmenes necesarios de reactivos para la eliminación de ADNg de muestras de ARN.

Componente	Vol. (μL) (1 rxn)
Buffer gDNA Wipeout, 7X	2
ARN (hasta 1 μg)	X (mín. 2)
Agua estéril libre de ARNasas	14 - X
Total	14 μL

Tabla 8. Volúmenes necesarios de reactivos para 1 reacción y para master-mix de 7 reacciones de retrotranscripción de ARN a ADNc.

Componente	Vol. (μL) (1 rxn)	Vol. (μL) (7 rxn)
Transcriptasa inversa	1	7
Buffer RT, 5X	4	28
RT Primer Mix	1	7
ARN sin ADNg	14	-
Total	20 μL	-

Después de incubar las muestras a 42°C durante 2 minutos en el termociclador, se preparó una master-mix (con los volúmenes de la tabla 8) con la transcriptasa inversa, buffer RT (que incluye Mg⁺⁺ y dNTPs) y RT primer mix, de los que se añadió 4 µL a los productos resultantes del paso anterior (14 µL de ARN sin ADNg). Una vez concluida la retrotranscripción a 42°C durante 15 minutos, fue necesario inactivar la transcriptasa inversa incubando las muestras a 95°C durante 3 minutos.

Tras cuantificar y analizar la pureza de las muestras de ADNc mediante el espectrofotómetro, se realizaron diluciones de 100 ng/µL para las 6 muestras que se almacenaron a 4°C, manteniendo los stocks originales a -20°C. Los cálculos se hicieron en base a la cuantificación previa del ARN, asumiendo una eficiencia del 100% de la retrotranscripción.

6.9. PCR cuantitativa (qPCR)

Las qPCR se realizaron con TB Green Premix Ex Taq II (Tli RNase H Plus) (Takara), que incluye todos los reactivos necesarios (a excepción de los *primers*) (tabla 9). Al principio se hicieron gradientes de temperatura de 55 a 65°C para determinar la temperatura de hibridación óptima y tras hallarla se siguieron llevando a cabo pruebas a esa temperatura (60°C). Uno de los experimentos se hizo con NZYSupreme qPCR Green Master Mix (2x) (NZYTech) para tratar de solventar problemas de contaminación de reactivos y se emplearon los volúmenes de la tabla 10. También se variaron las concentraciones de *primers* con el mismo objetivo.

Tabla 9. Volúmenes empleados para qPCR con TB Green Premix Ex Taq II (Tli RNase H Plus) (Takara).

Componente	Vol. (1 rxn) (µL)	Concentración final
TB Green Premix Ex Taq (Tli RNase H Plus) (2X)	12,5	1X
Primer Fw (10 µM)	0,5	0,2 µM
Primer Rev (10 µM)	0,5	0,2 µM
cDNA (100 ng/µL)	1	100 ng
Agua libre de nucleasas	10,5	
Total	25	

Tabla 10. Volúmenes empleados para qPCR con NZYSupreme qPCR Green Master Mix.

Componente	Vol. (1 rxn) (µL)	Concentración final
NZYSupreme qPCR Green Master Mix (2X)	10	1X
Primer Fw (10 µM)	0,8	400 nM
Primer Rev (10 µM)	0,8	400 nM
cDNA (100 ng/µL)	1	100 ng
Agua libre de nucleasas	7,4	
Total	20	

Estas dificultades dadas por la aparente contaminación de los blancos nos llevaron a preparar las reacciones en llama y posteriormente, en campana de flujo laminar. Ya que el problema no cesaba, se hizo una qPCR con los *primers dmrt1-13* y los *primers eEf1a1* (para el gen *housekeeping* del mismo nombre, que codifica para el factor de elongación eucariota 1-alpha) a modo de control positivo, para ADNg y ADNc.

Las reacciones de amplificación fueron llevadas a cabo en el equipo CFX Connect Real-Time PCR Detection System (Bio-Rad) a través del software Bio-Rad CFX Manager. En todos los casos se utilizó el programa de amplificación de la tabla 11, capturando imágenes de la placa tras terminar cada ciclo y al final de la curva de *melting*.

Paso		Temperatura	Tiempo
Desnaturalización previa		95°C	30 segundos
Desnaturalización	40 ciclos	95°C	5 segundos
Hibridación/anillamiento y extensión		60°C	30 segundos
Curva de <i>melting</i>		65 a 95°C	5 segundos para cada incremento de 0,5°C

Tabla 11. Condiciones de tiempo y temperatura para qPCR.

Después de finalizar la amplificación, si los resultados dados por las temperaturas de *melting* no eran suficientes para esclarecer la naturaleza del producto obtenido, se transfería 15 µL de cada pocillo de la microplaca a tubos Eppendorf para visualizar los productos en geles de agarosa y tener certeza de que fueran únicos.

6.10. Alineamiento de la secuencia del gen *dmrt1* con el genoma

Las secuencias utilizadas en el diseño de *primers* de los exones e intrones del gen *dmrt1* para *S. senegalensis* corresponden a las obtenidas por Cross et al. [35], que para ello había alineado la secuencia de ADNc obtenida con el clon BAC que contiene el gen *dmrt1* (BAC 48K7). La secuencia completa del gen obtenida tras agrupar las secuencias exónicas e intrónicas se guardó en un archivo FASTA para su uso posterior.

Durante la realización de este trabajo, se publicó otro estudio en el que se había generado un ensamblaje *de novo* para un lenguado macho, que tras combinarse con un mapa de SNP de alta densidad y ser mejorado a partir de lecturas de secuenciación dio a lugar un borrador del genoma de la especie de 609 Mb y con 21 pseudo-cromosomas [44]. Para comparar la estructura genómica de *dmrt1* con este borrador se hizo uso de BLAST+ (v. 2.9.0-2), una herramienta que permite ejecutar BLAST de manera local a través de la línea de comandos de Linux. La base de datos fue creada con el comando `makeblastdb -in sse.fasta -dbtype nucl -out sse`, que indexa el ensamblaje del genoma de macho (archivo “sse.FASTA”; número de acceso de GenBank GCA_019176455.1) para que pueda ser consultado de forma eficiente [48]. Luego, para el análisis BLAST de nucleótidos se ejecutó el comando `blastn -db sse -query dmrt1.fasta -out resultados.txt -outfmt 6 -evalue 0.001`, donde “*dmrt1.FASTA*” se corresponde a la secuencia completa del gen. Los resultados se exportaron a una hoja de cálculo y se seleccionaron aquellos alineamientos con mayor *bitscore* y un *e-value* de 0, que pertenecían todos a un mismo código de acceso (CM032858.1, con un total de 11 alineamientos). Luego estos se ordenaron de menor a mayor posición de inicio en la secuencia *query* (o problema) y se trató de reconstruir una serie de alineamientos que cubriera la totalidad de la secuencia del gen con solapamiento mínimo y que también estuviera ordenado de menor a mayor posición de inicio en la secuencia *subject* (o sujeto).

Más tarde se verificó el alineamiento utilizando la plataforma online de BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>), una vez se había incluido la base de datos anteriormente empleada (GCA_019176455.1) en la herramienta, seleccionando como programa `blastn` y dejando el resto de parámetros por defecto [49].

7. RESULTADOS Y DISCUSIÓN

7.1. Diseño de *primers*

En un primer momento, el objetivo era diseñar dos parejas de *primers* que mediante amplificación larga cubrieran la totalidad de la secuencia del gen para secuenciar varios individuos machos y hembras. Sin embargo, debido a dificultades posteriores, se tuvo que buscar otras dos parejas alternativas. Finalmente, ante la falta de resultados positivos y para continuar caracterizando el gen y buscar una posible relación con la determinación del sexo, se optó por diseñar una pareja de *primers* que permitiera la amplificación del intrón 3 a fin de una búsqueda de polimorfismos mediante análisis por digestión con enzimas de restricción y otras dos parejas dentro de los exones I y IV con el objeto de realizar una cuantificación de la expresión del gen por qPCR.

El motivo detrás de la elección del intrón 3 para la digestión radica en que este se encuentra delimitado por los exones III y II', que debido al patrón específico de duplicación ocurrido en *dmrt1* permite el diseño de *primers* específicos, como no ocurre en el resto de exones, que se encuentran duplicados y por tanto son capaces de dar más de un producto PCR. Los exones I y IV, por su parte, fueron escogidos por estar presentes en prácticamente todos los distintos transcritos de ARNm descritos hasta la fecha, lo cual permitiría una cuantificación total de la expresión del gen (además, la duplicación del exón I resulta en el exón no funcional $\Delta I'$, que al carecer de ORF no podría ser transcrito a ARNm, por lo que la duplicación tampoco debería acarrear problemas de inespecificidad para esta pareja de *primers*). En la tabla 12 se presentan estas 7 parejas de *primers*, junto con su región de estudio en *dmrt1*, su finalidad en el trabajo y el tamaño de producto PCR esperado.

Tabla 12. Secuencia, región amplificada y tamaño esperado para *primers*. F = forward, R = reverse.

Primer	Secuencia dirección 5'-3'	Región a estudiar	Finalidad	Tamaño del producto PCR esperado (pb)
<i>dmrt1</i> -F10	AAACAGGTGCCGGACTTCAC	Entre exones I-IV	Amplificación larga	15.470
<i>dmrt1</i> -R10	TTGTTATTGCTGCCCTCCAG			
<i>dmrt1</i> -F11	GGAGGGCAGCAATAACAACA	Entre exones IV-V	Amplificación larga	15.119
<i>dmrt1</i> -R11	ATCTCACACTCGGCCTTGAC			
<i>dmrt1</i> -F12	ACGGCAGTCTCTACAACCTACC	Intrón 3 (entre exones III-II')	Análisis por digestión con enzimas de restricción	2.522
<i>dmrt1</i> -R12	CTGCAAAATACCAAGCTCCTCC			
<i>dmrt1</i> -F13	CACGGCTACGTGTCTCCACT	Exón I	qPCR	110
<i>dmrt1</i> -R13	TGTGCAGCCATGACTCTCTG			
<i>dmrt1</i> -F14	TCTCTTGGTGACGGCTGC	Exón IV	qPCR	196
<i>dmrt1</i> -R14	CCATGGTCTCAGATAGTTGTAG			
<i>dmrt1</i> -F15	GCAAGGACAAGCAGAGCAAACAGG	Entre exones I-IV	Amplificación larga	15.418
<i>dmrt1</i> -R15	TAGGTGGCGGCTGGGTAGAAGG			
<i>dmrt1</i> -F16	AGCCACAACATGTCCTCCC	Entre exones IV-V	Amplificación larga	15.233
<i>dmrt1</i> -R16	GAAGCTGATCTCACACTCGG			

En la tabla 13 se muestran las características de cada *primer* dadas por Primer3. *A priori*, parecen *primers* adecuados puesto que la mayoría presentan valores deseables de longitud, Tm, contenido en GC y complementariedad (tanto la autocomplementariedad como la complementariedad entre ambos miembros de cada pareja, dado en general y para el extremo 3').

Tabla 13. Parámetros de los *primers* diseñados usando Primer3 para el gen *dmrt1* en *S. senegalensis*.

Primer	Longitud (pb)	Tm (°C)	%GC	GC en 3'	Autocompl. (any)	Autocompl. en 3' (3')	Compl. (pair any)	Compl. en 3' (pair 3')
<i>dmrt1</i> -F10	20	62,43	55	2	4	2	4	1
<i>dmrt1</i> -R10	20	61,16	50	3	3	2		
<i>dmrt1</i> -F11	20	61,03	50	2	3	0	3	0
<i>dmrt1</i> -R11	20	60,27	55	2	4	1		
<i>dmrt1</i> -F12	21	59,18	52,38	3	0	0	0	0
<i>dmrt1</i> -R12	21	58,89	52,38	4	0	0		
<i>dmrt1</i> -F13	20	61,33	60	3	4	1	5	2
<i>dmrt1</i> -R13	20	61,21	55	3	6	1		
<i>dmrt1</i> -F14	18	59,66	61,11	4	0	0	0	0
<i>dmrt1</i> -R14	23	57,79	47,83	2	2,57	0		
<i>dmrt1</i> -F15	24	58,96	54,17	3	2	0	3	0
<i>dmrt1</i> -R15	22	59,72	63,64	3	3	0		
<i>dmrt1</i> -F16	19	59,62	57,89	4	0	0	0	0
<i>dmrt1</i> -R16	20	58,08	55	4	0	0		

Los resultados del análisis con OligoAnalyzer realizado después de probar los *primers* se recogen en la tabla 14. En ella se muestra el número total de homodímeros, heterodímeros y horquillas capaces de formar cada uno de los *primers*, indicando asimismo la energía libre de Gibbs (ΔG) de la estructura más estable (con el valor más negativo, y por tanto, la más indeseable al tener una mayor probabilidad de darse en el medio de reacción). También se muestra la Tm de las horquillas más estables, lo cual es importante puesto que es preferible tener valores más bajos, ya que entonces se requiere menos energía para romper los enlaces por puente de hidrógeno. Además, si alguna de las horquillas tuviera una Tm por encima de la temperatura de hibridación, podría ser capaz de impedir la hibridación de los *primers* con el molde. Para estos *primers*, ninguna de las horquillas que se pueden formar debería implicar un riesgo significativo en las reacciones de PCR, dado que tienen valores de estabilidad bajos (todas las ΔG son altas y positivas) y que sus Tm están por debajo de las temperaturas de hibridación de los *primers*. Una excepción notable se trata de una de las horquillas formada por el *primer dmrt1*-F13, representada en la figura 11, que posee una estabilidad superior a las demás (0,03 kcal/mol) y cuya Tm (59,6°C) es muy cercana a la temperatura de hibridación (60°C), por lo que hay una pequeña probabilidad de que afecte negativamente a la amplificación.

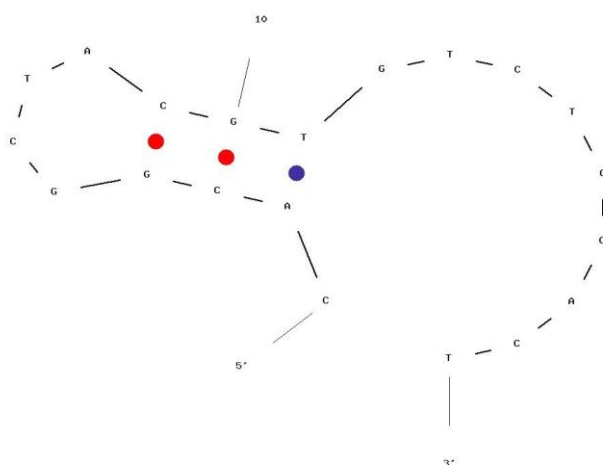


Figura 11. Horquilla formada en el *primer dmrt1*-F13. Fuente: OligoAnalyzer.

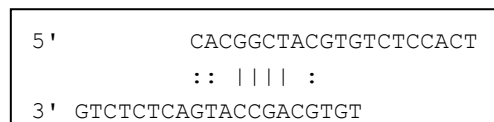


Figura 12. Heterodímero formado entre los *primers dmrt1*-F13 y R13.

Aun así, se da en una posición alejada del extremo 3' y solo tiene tres puntos de unión, por lo que no debería interferir en la reacción [40]. Con respecto a los heterodímeros, aunque ninguna pareja de *primers* forma un dímero con una estabilidad excesiva, al ser todos los valores negativos no se puede descartar que produzcan estructuras relativamente estables potencialmente problemáticas durante las PCR. Entre ellas destaca la de la pareja 13, ilustrada en la figura 12, que aunque tiene cuatro puntos de unión, al no ser próxima al extremo 3' se minimizan los posibles efectos adversos sobre la amplificación (ya que de modo contrario, podría impedir la acción de la polimerasa, que se une a 3'). Finalmente, en vista de los valores de estabilidad obtenidos, todas las parejas de *primers* son capaces de formar homodímeros estables, hecho que podría tener efectos adversos en los experimentos posteriores.

Tabla 14. Formación de dímeros y horquillas a partir de las parejas de *primers* con sus valores de estabilidad (ΔG), resultado del análisis con OligoAnalyzer. HMD: homodímero, HTD: heterodímero. Valor límite recomendado por el programa: no menor a -9 kcal/mol.

<i>Primers</i>	Nº de HMD	ΔG del HMD más estable (kcal/mol)	Nº de horquillas	ΔG de la horquilla más estable (kcal/mol)	Tm de la horquilla más estable (°C)	Nº de HTD	ΔG del HTD más estable (kcal/mol)
<i>dmrt1-F10</i>	10	-9,75	8	2,42	-30,4	19	-5,24
<i>dmrt1-R10</i>	9	-3,55	1	1,87	41,4		
<i>dmrt1-F11</i>	5	-3,14	0	-	-	16	-6,21
<i>dmrt1-R11</i>	6	-9,28	6	2,63	5,8		
<i>dmrt1-F12</i>	10	-3,61	6	2,28	5,8	11	-6,69
<i>dmrt1-R12</i>	7	-7,05	3	1,82	33,4		
<i>dmrt1-F13</i>	8	-6,91	2	0,03	59,6	13	-7,81
<i>dmrt1-R13</i>	14	-7,05	4	1,52	34,6		
<i>dmrt1-F14</i>	7	-3,61	4	1,63	-0,8	15	-5,02
<i>dmrt1-R14</i>	13	-11,52	15	2,09	13,6		
<i>dmrt1-F15</i>	6	-3,14	3	2,41	33,9	17	-4,74
<i>dmrt1-R15</i>	9	-3,61	5	2,41	29,7		
<i>dmrt1-F16</i>	5	-8,07	2	1,36	14,0	16	-4,74
<i>dmrt1-R16</i>	10	-6,34	3	1,50	-21,7		

En cuanto al análisis de los amplicones de los *primers dmrt1-13* y *14* con mFOLD, se observan estructuras secundarias muy inestables (teniendo las más estables unas ΔG de -0,85 y -3,60 kcal/mol, respectivamente) que son las representadas en la figura 13. En teoría, no deberían presentar ningún problema grave porque no interfieren con el extremo 3' del producto (que debe quedar libre para asegurar el correcto funcionamiento de la polimerasa) y tienen valores de energía libre no demasiado bajos. Bien es cierto que hay que prestar atención al hecho de que sus Tm, 63,5 y 67,2°C, están por encima de la temperatura de hibridación de los *primers* (60°C), por lo que podrían no desdoblarse del todo durante la reacción e impedir el avance de la polimerasa a lo largo del molde. Sea como fuere, no son resultados concluyentes y habría que probar los *primers in vitro*.

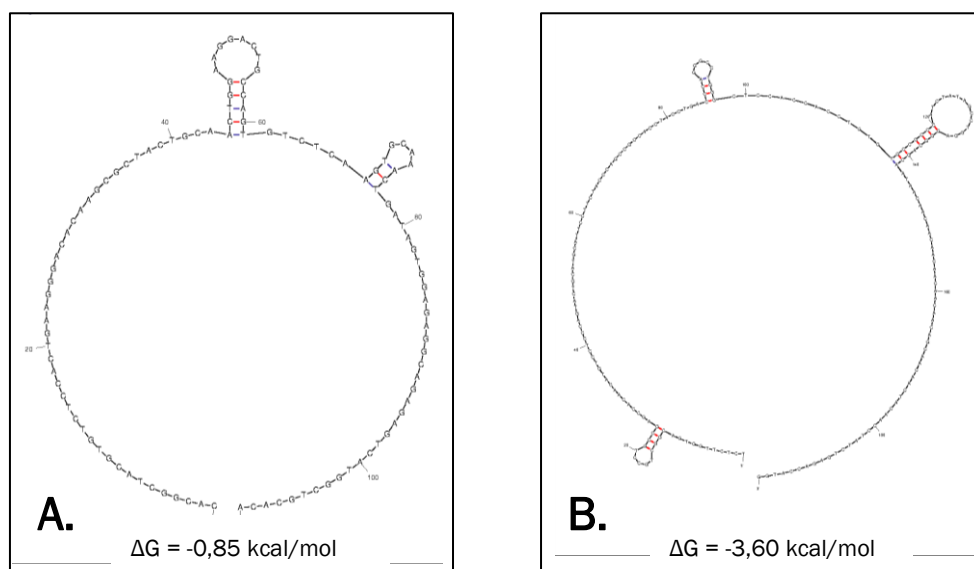


Figura 13. Formación de estructuras secundarias inestables por los amplicones de las parejas de *primers dmrt1-13* (A) y *dmrt1-14* (B). Fuente: mFOLD.

7.2. PCR *in silico*

Como ya se explicó anteriormente, no fue posible comprobar la especificidad de las parejas de *primers* de forma teórica hasta después de finalizar los experimentos, ya que hasta entonces no estaba disponible la secuencia del genoma de la especie. El uso de iPCReSS resultó en archivos de texto para cada pareja de *primers* en los que se listaban todos los posibles productos (no sólo de la hibridación de los *primers* Fw-Rev con el ADN, que es lo esperado; sino también entre ellos mismos, Fw-Fw y Rev-Rev, y en sentido contrario, Rev-Fw) con el número de emparejamientos incorrectos (*mismatches*) que deben darse para cada uno. En la tabla 15 se muestra un resumen de los resultados.

Tabla 15. Productos resultantes de la PCR *in silico* con la secuencia del genoma del lenguado y las diferentes parejas de *primers*. X: producto para el que fueron diseñados (coincidencia perfecta, con 0 *mismatches*). Si no se indica expresamente, presentan 4 emparejamientos incorrectos (*mismatches*) para cada *primer*.

<i>Primers</i>	Productos generados con <i>primers</i> X-X (pb)			
	Fw-Rev	Rev-Fw	Fw-Fw	Rev-Rev
<i>dmrt1-10</i>	292, 847, 14.545, 15.500 , 17.598	10.865, 16.789, 14.436	3.273, 4.865, 8.233, 18.454, 19.724	8 productos entre 10.149 y 19.597
<i>dmrt1-11</i>	5.629, 9.369, 15.036	134, 147, 6.131, 13.747, 14.121, 18.879, 18.892	14 productos entre 330 y 19.576	6.885, 11.976
<i>dmrt1-12</i>	2.514	–	–	–
<i>dmrt1-13</i>	110	–	–	2.920
<i>dmrt1-14</i>	196	–	27 productos entre 109 y 2.864	–
<i>dmrt1-15</i>	15.448	–	–	–
<i>dmrt1-16</i>	15.150 y 10 productos entre 1.506 y 17.203	14 productos entre 1.113 y 16.727	58 productos entre 1.436 y 19.358	1.404, 5.045 (con 4 y 0 <i>mismatches</i>), 5.498, 13.110, 13.777, 16.327

Algunos de los productos esperados (exceptuando los de los *primers dmrt1-13 y 14*) difieren en varias decenas de pares de bases al tamaño para el que fueron inicialmente diseñados, ya que para el diseño se empleó la secuencia del gen obtenida por alineamiento entre el ADNc y el clon BAC al que pertenece, mientras que para la PCR *in silico* se tomó el genoma de un individuo distinto como referencia. Aunque el análisis BLAST anterior mostró una altísima similitud entre la secuencia con la que se trabajó y el genoma ensamblado, es razonable que el alineamiento no sea exacto al 100% al tratarse de distintos individuos y métodos de secuenciación, más aún considerando que las diferencias se dan en regiones intrónicas (de alta variabilidad).

En base al número de productos PCR, los *primers dmrt1-12, 13 y 15* parecen ser altamente específicos, pero los *primers dmrt1-10, 11, 14 y 16* presentan un número potencialmente elevado de productos inespecíficos, lo que podría llevar a que la amplificación no sea única, como se confirmará en los siguientes apartados (especialmente los *primers 16*, que cuentan con un número excepcional de productos de la hibridación con Fw-Fw). Bien es cierto que la inmensa mayoría de estos productos sólo son posibles permitiendo 4 emparejamientos incorrectos en cada *primer* (es decir, el *primer* con sentido, que de media tendrá 20 pb, hibrida con el ADN en todas sus posiciones menos en 4; y lo mismo para el *primer* antisentido, que sólo hibridará en 16 posiciones) lo cual afecta a la estabilidad del dúplex *primer*-molde y consecuentemente, a la eficiencia de la extensión por parte de la polimerasa.

Por eso, la pregunta de si la hibridación será específica o no debe ser resuelta en el laboratorio. Una excepción digna de mencionar es el producto de 5.045 pb fruto de la hibridación con ambos *primers* Rev de la pareja 16, dado que el *primer* que actúa como antisentido se une perfectamente al molde (mientras que el con sentido presenta 4 *mismatches*), así que es probable que se dé *in vitro*.

Además, pese a que cualquier emparejamiento incorrecto puede afectar a la especificidad de la PCR, son cuando se dan en el extremo 3' de cualquiera de los *primers* cuando la extensión puede verse perjudicada y no darse la amplificación de ese producto (lo que sería beneficioso en nuestro caso, ya que el objetivo es que los *primers* den un único producto PCR específico), con lo que los productos que teóricamente se formarían bajo esas condiciones no deberían suponer un problema real en la amplificación [50].

Un último punto a destacar es que aunque *in silico* se obtengan productos de gran tamaño (más de 10.000 pb), en la realidad es complicado conseguirlos debido a la formación de estructuras secundarias a medida que se va amplificando el producto y a otros factores durante la amplificación, como los tiempos de extensión y las enzimas usadas, además de la calidad del ADNg. Así pues, los más preocupantes serían los productos de menor tamaño, ya que hay mucha más probabilidad de que se formen y de que pongan en riesgo la especificidad de la reacción.

7.3. Extracción de ADN genómico

Por electroforesis en gel de agarosa se comprobó la integridad del ADN_g extraído, dando como resultado una banda alrededor de las 23.130 pb (banda más pesada del marcador molecular usado) para cada una de las muestras y un ligero *smear* por debajo de las bandas, que podría deberse a contaminación por proteínas o a degradación parcial del ADN. Los resultados de la cuantificación y determinación de la pureza de las seis muestras de ADN_g se presentan en la tabla 16. La mayoría de ellas tienen un cociente A260/A280 cerca de 1,8, que es el valor adecuado para el ADN puro, por lo que no parece observarse una contaminación proteica excesiva. El otro cociente de absorbancias, A260/A230, debería tener valores entre 2-2,2 si estuviéramos tratando con muestras de ácidos nucleicos puros. Sin embargo, todas las muestras se alejan bastante de este valor de referencia (excepto el ADN_g de la hembra 2) y muestran valores considerablemente menores, que probablemente se deba a contaminación por carbohidratos y por compuestos usados en la extracción, como cloroformo.

Con respecto a la cuantificación, destaca la concentración del ADN_g de la hembra 2, que es muy inferior al del resto. Este hecho, junto a los valores de absorbancia de la muestra (que denotan una elevada pureza), implicaría que la extracción para esta muestra ha sido exitosa, mientras que para el resto de muestras el aparente rendimiento (viendo las concentraciones) sería mayor al real, por la presencia de otros compuestos contaminantes. No es inesperado que la pureza sea tan baja, puesto que el protocolo de extracción utilizado fue rápido y no repetía ciclos de precipitación y lavado del ADN.

De cualquier forma, las muestras tenían la suficiente pureza y concentración como para seguir trabajando con ellas en posteriores experimentos.

Tabla 16. Resultados de la cuantificación y determinación de la pureza de las muestras de ADN genómico extraídas. A260/280 para ADN puro: 1,8. A260/230 para ADN y ARN puro: 2-2,2.

Muestra	Concentración (ng/μL)	A260/A280	A260/A230
Macho 1	781,2	1,77	1,3
Macho 2	1477,1	1,74	0,86
Macho 3	1087,9	1,81	1,08
Hembra 1	1035,7	1,93	1,47
Hembra 2	314,9	1,82	1,91
Hembra 3	1082,1	1,78	1,01

7.4. Amplificación por PCR de fragmentos largos (con *primers dmrt1-10, 11, 15 y 16*)

Las múltiples pruebas realizadas con las cuatro parejas de *primers* destinadas a amplificar la totalidad del gen *dmrt1* (*primers dmrt1-10 y 15* para la primera mitad; y *11 y 16* para la segunda mitad) no resultaron en bandas correspondientes a los tamaños esperados en ninguno de los casos (que era de alrededor de 15.000 pb para cada una de las parejas de *primers*). Tampoco se consiguió amplificar los fragmentos deseados alargando los tiempos de extensión (como es sugerido por el proveedor de la polimerasa Q5 para amplicones especialmente largos), ni variando la temperatura de extensión y tampoco empleando buffer Enhancer GC diseñado para productos con alto contenido en GC.

Los *primers* 11 no produjeron ningún tamaño de banda, mientras que la pareja 10 dio una gran cantidad de amplificación inespecífica con productos que iban desde las ~8.000 hasta los cientos de pares de bases. Especialmente llamativa fue una banda de unas 1.800 pb que aparecía en todas las muestras (incluyendo un control positivo, correspondiente a un individuo proporcionado por un compañero, que en una PCR inicial también había dado amplificación inespecífica) a tamaños ligeramente distintos, como se puede comprobar en el gel de la figura 14. Tanto fue así que se decidió cortar algunas de estas bandas de 1.800 pb para ser secuenciadas, como se desarrollará en uno de los próximos apartados. Cabe mencionar que las únicas bandas que podrían corresponderse con los tamaños predichos por la PCR *in silico* son las observadas en todas las muestras (excepto en el control positivo) entre las 9.416 y 6.557 pb (que podría ser un producto de 8.233 pb) y por debajo de las 4.361 pb (que se trataría de un producto de 3.273 pb). Si bien las uniones inespecíficas de los *primers* podrían verse parcial o completamente explicadas por un exceso de ADN o una temperatura de hibridación demasiado baja, esto no da respuesta a por qué no fueron obtenidas en la PCR *in silico*, teniendo en cuenta que se permitieron hasta 4 emparejamientos incorrectos. También se desconoce el motivo por el cual la amplificación con el individuo del control positivo no produjo bandas inespecíficas como el resto de muestras (más allá de la de 1.800 pb), pero podría deberse al simple hecho de que se trata de un individuo completamente distinto que podría presentar variaciones en el genoma que no den lugar a estos productos. Otra explicación podría ser diferencias en la cantidad inicial de ADN en las reacciones.

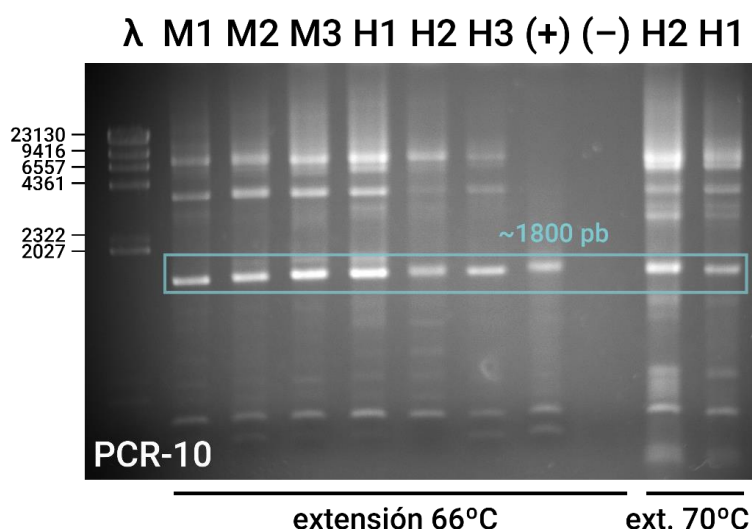


Figura 14. Amplificación de ADN_g con Q5 y los *primers* *dmrt1-10* a distintas temperaturas de extensión (66 y 70°C) para diferentes individuos. λ: Lambda DNA/HindIII. M1-H3: ADN_g de tres individuos macho y tres individuos hembra. (+): individuo que previamente había dado amplificación inespecífica. (-): blanco/control negativo. H2: hembra 2. H1: hembra 1.

En general, la presencia de bandas inespecíficas (que además no pudieran ser predichas con la PCR *in silico* realizada *a posteriori*) es un fenómeno recurrente a lo largo del trabajo que podría estar relacionado en cierta medida con el alto grado de duplicación de genes que presentan los peces teleósteos, como resultado de un evento de poliploidía que tuvo lugar durante la evolución de este clado [51].

Es posible que la existencia de polimorfismos en estas regiones duplicadas, y más generalmente la variación genética, fuera responsable de la incapacidad de la PCR *in silico* de detectar el gran número de productos alternativos amplificados con la mayoría de parejas de *primers*

Aunque se podría haber tratado de reducir la inespecificidad de las amplificaciones modificando ciertos factores (como el número de ciclos, la concentración de reactivos como *primers*, MgCl₂ y ADN, y la temperatura de hibridación) hasta optimizar las condiciones de PCR, no se consideró oportuno puesto que en ninguno de los casos se había obtenido el producto PCR deseado.

Al igual que con la pareja anterior, las parejas *dmrt1*-15 y 16 dieron una amplificación inespecífica que no pudo ser explicada por la PCR *in silico*, dando además patrones de bandas más variables que en el caso anterior, como se observa en la figura 15, sin rastro de los productos esperados de ~15.000 pb.

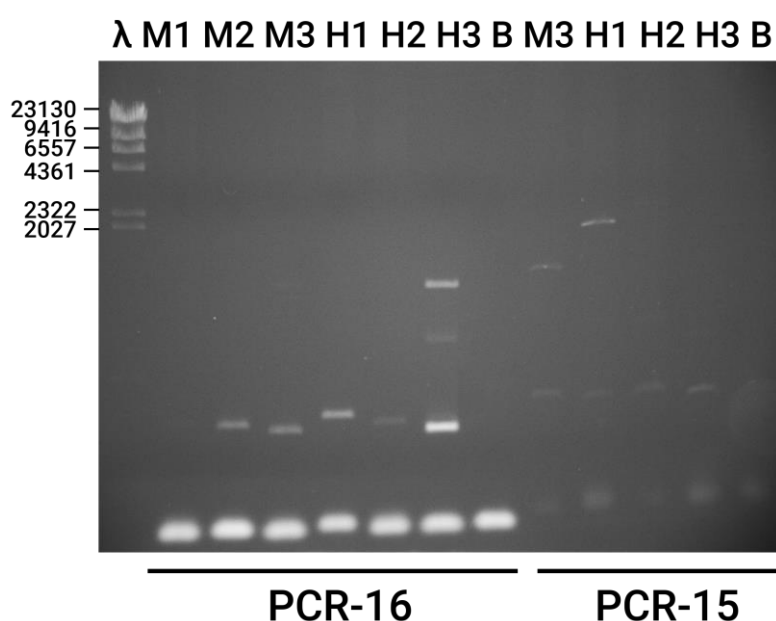


Figura 15. Amplificación de ADNg con Q5 y los *primers dmrt1*-16 y 15 para diferentes individuos. λ: Lambda DNA/HindIII. M1-H3: ADNg de tres individuos macho y tres individuos hembra. B: blanco/control negativo.

7.5. Amplificación por PCR del intrón 3 (con *primers dmrt1*-12)

La primera prueba con los *primers dmrt1*-12 con la polimerasa Q5 dio como resultado distintos patrones de bandas inespecíficas para los seis individuos, además del producto esperado de unas 2.500 pb (correspondiente a la totalidad del intrón 3 de *dmrt1*). Al no ser un fragmento excesivamente largo, se trató también de amplificar con la polimerasa NZYTaQ II, obteniéndose una única banda tenue del tamaño esperado para tres de los individuos y ninguna banda prominente para los otros tres (excepto de dímeros de *primers*, igual de intensas en todas las muestras, incluida el blanco). Ya que una de las posibles causas para la aparición de bandas inespecíficas es que la temperatura de hibridación sea demasiado baja, se probó incrementándola en 3°C para la amplificación con Q5, sin observarse mejora alguna; al contrario, seguía habiendo bandas inespecíficas y no se obtenía el producto PCR.

El uso de buffer Enhancer GC para altos contenidos en GC tampoco aportó resultados: desaparecían tanto las bandas inespecíficas como el producto, observándose únicamente dímeros de *primers*. Como siguiente estrategia se optó por realizar un gradiente de temperaturas de hibridación con Q5, de 60 a 72°C (anteriormente se venía usando 67°C) y se determinó la temperatura óptima, de 71.1°C, en la que se minimizaba el número de bandas inespecíficas y aparecía el producto deseado (figura 16).



Figura 16. Gradiente de temperaturas de hibridación para amplificación de ADN_g con Q5 de un individuo con *primers dmrt1-12*. λ: Lambda DNA/HindIII. HII: HyperLadder II. A: 72°C. B: 71,1°C. C: 69,6°C. D: 67,2°C. E: 64,5°C. F: 62,4°C. G: 60,9°C. H: 60°C.

Se probó a usar las nuevas condiciones de temperatura para el resto de individuos, obteniéndose en general buenos resultados de amplificación (bandas únicas y pequeñas bandas inespecíficas tenues para algunos de los individuos), si bien los blancos parecían estar contaminados, pues aparecían algunas bandas (de diferentes tamaños a las observadas en las muestras).

Después de numerosas pruebas destinadas a solventar el problema de la contaminación, no fue posible replicar las amplificaciones previas y apenas se consiguió producto PCR del tamaño esperado para la mitad de las muestras. En una prueba posterior cuyo objetivo era determinar el origen de la contaminación, se vio que la polimerasa NZYtaq II (que había sido descartada en experimentos previos en favor de Q5) daba una amplificación extremadamente limpia del fragmento, sin aparición de bandas inespecíficas, empleando las condiciones de tiempo y temperatura de Q5. Al repetir el experimento para el resto de individuos (figura 17), se obtuvieron bandas únicas de producto PCR de ~2.500 pb (aunque en presencia de dímeros de *primers*). Por motivos desconocidos, no se pudieron replicar estas últimas bandas con la polimerasa Taq: o bien la amplificación fallaba por completo o sólo se observaban dímeros de *primers*.

Tampoco dio resultados el uso de diferentes concentraciones de *primers* ni el reemplazo de los reactivos de PCR (incluyendo agua, Taq, MgCl₂, buffer, todos varias veces; ADNg, preparando nuevas alícuotas; e incluso *primers* que fueron pedidos de nuevo). Después de múltiples intentos, utilizando nuevas alícuotas de polimerasa, MgCl₂ y buffer se consiguieron bandas similares a las obtenidas en la figura 20 que finalmente pudieron ser utilizadas para la digestión con enzimas de restricción, cuyos resultados se presentan más adelante.

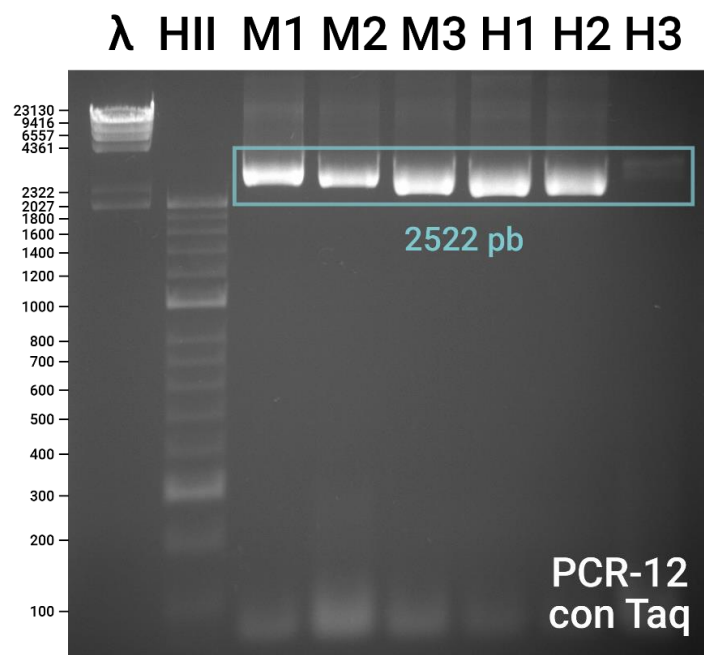


Figura 17. Amplificación de ADNg con NZYTaq II y con *primers dmrt1-12*. λ: Lambda DNA/HindIII. HII: HyperLadder II. M1-H3: ADNg de tres individuos macho y tres individuos hembra.

7.6. Amplificación por PCR de fragmentos cortos (con *primers dmrt1-13* y 14)

Antes de usar las parejas de *primers dmrt1-13* y 14 en PCR cuantitativa, se hicieron pruebas de amplificación estándar (con NZYTaq II), tanto con ADNc como ADNg, para confirmar que se obtenían los productos correctos (de 110 y 196 pb, respectivamente). Cabe destacar que si cada pareja de *primers* hubiera sido diseñada para amplificar varios exones, los tamaños de los fragmentos tras amplificar con ADNg no concordarían al presentar además los intrones. Sin embargo, como en este caso se trataba de amplificar los exones I y IV por separado, cada uno con su propia pareja de *primers*, se esperaba obtener los mismos tamaños.

Para los *primers 13* se usó una temperatura de hibridación de 57,5°C a lo largo de las amplificaciones estándar. Debido a una confusión en el diseño preliminar de *primers*, en un principio se pensó que el producto PCR debería ser de 219 pb (en lugar de 110) por lo que las pruebas con ADNg (figura 18A) parecieron tener éxito y se continuó trabajando con ellos. El hecho de que la banda de ~200 pb no apareciera en el blanco (del que se desconoce el motivo; en todas las demás amplificaciones, incluyendo la mostrada en la figura 18B, aparecían tres bandas relativamente pequeñas, tanto en las

muestras como en los blancos) y de que esa misma banda tuviera un tamaño ligeramente distinto para cada individuo (que podría deberse a que la agarosa no se disolviera por completo, dando a una migración irregular de los fragmentos) llevó a pensar que la amplificación fue correcta, cuando realmente no lo fue al no haberse obtenido el tamaño esperado. En la figura 18B se muestra la misma PCR utilizando muestras de ADNc en su lugar: aparecen bandas difusas de 50, 100 y 200 pb en todos los pocillos, incluido el blanco.

Ya que los dímeros de *primers* suelen tener menores tamaños (en torno a 50 pb) se sospechó inicialmente de una posible contaminación. Ante este problema se cambiaron los reactivos varias veces (agua, Taq, $MgCl_2$ y buffer), sin observar ningún cambio. Teniendo en cuenta que los resultados del análisis de dimerización de los *primers* (homodímeros y heterodímeros) fueron particularmente malos para esta pareja, se cree que los distintos dímeros de *primers* podrían estar uniéndose entre sí y aparentando un mayor tamaño, dando por tanto bandas tan improbables como las observadas.

Por otro lado, no parece obtenerse el producto deseado; además, si se obtuviera sería difícil de distinguirlo de las bandas de dímeros de *primers*, que son de un tamaño similar. Es verdad que para la muestra 6 de la figura 18B la banda por encima de 100 pb (que podría ser nuestro producto) tiene una intensidad mayor a las mismas bandas de otros pocillos, mientras que la intensidad de la banda inferior por encima de 50 pb se ve disminuida (y la banda justo por debajo de 200 pb, inexistente) lo que sugiere que la amplificación correcta del exón podría estar compitiendo con la formación de dímeros de *primers*. En cualquier caso, las bandas de “dímeros de *primers*” impiden confirmar la existencia del amplicón, por lo que esta pareja de *primers* debería haber sido descartada antes de seguir con el paso de qPCR (y no se hizo, como ya se ha mencionado, debido a un error de cálculo).

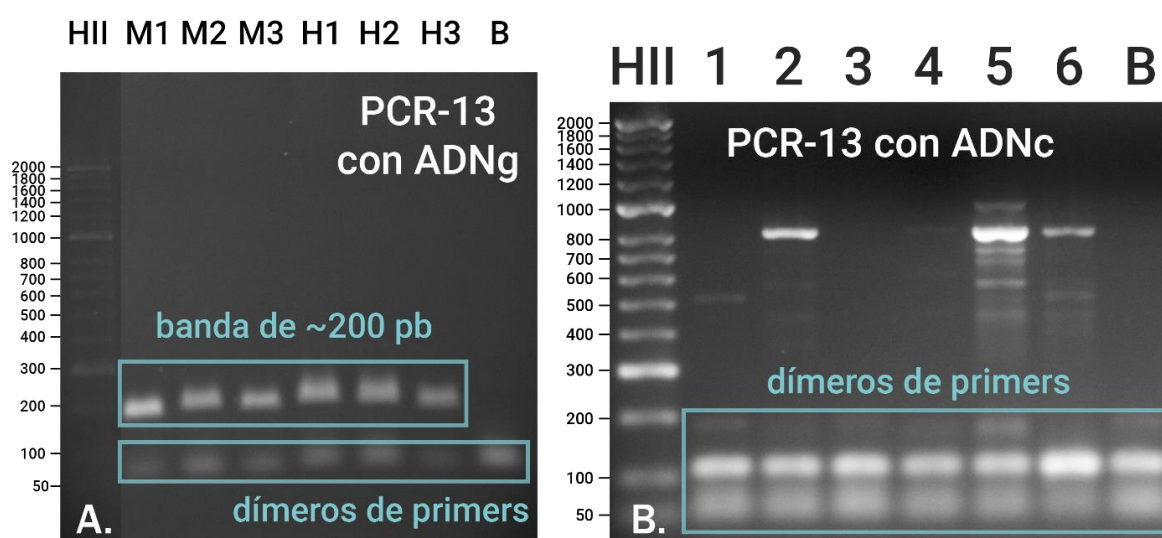


Figura 18. A: amplificación de ADNg con NZYTaq II y con *primers dmrt1-13*. HII: HyperLadder II. M1-H3: ADNg de tres individuos macho y tres individuos hembra. B: blanco/control negativo. B: amplificación de ADNc con NZYTaq II con *primers dmrt1-13*. HII: HyperLadder II. 1-6: ADNc de macho inmaduro WT, hembra inmadura WT, macho inmaduro F1, hembra inmadura F1, macho maduro F1, hembra madura F1. B: blanco/control negativo.

También llama la atención la presencia de bandas inespecíficas en algunas de las muestras, particularmente la 2, 5 y 6 donde se observan bandas brillantes de 800 pb que de nuevo no pueden ser explicadas por la PCR *in silico*. El único producto indeseado que podría esperarse a partir de la secuencia del genoma del lenguado tiene un tamaño de 2.920 pb, formado con los *primers* 13 Rev-Rev.

Una posible explicación es que el ARN a partir del cual se obtuvo el ADNc estuviera contaminado en cierta medida con ADNg (que no es el mismo que las muestras de ADNg con las que se hizo la otra amplificación; los individuos de los que se extrajeron el ADNg y ADNc eran completamente distintos entre sí, por lo que los resultados no tienen que concordar necesariamente con el otro gel careciente de bandas inespecíficas). De todos modos, esto no aclararía la naturaleza de los productos inespecíficos (porque la PCR *in silico* con ADNg no dio tamaños similares) así que se podría pensar que el ARNm, al haberse eliminado los intrones, podría tener nuevas dianas (no existentes en el ADNg) con las que podrían hibridar los *primers*. En dos de las muestras (1 y 6) se observó una tenue banda por encima de las 500 pb que podría corresponderse con el ARNm entre el exón I y la duplicación parcial del mismo exón (teniendo en cuenta que el *primer* Rev hibrida en su totalidad con parte de la secuencia repetida, presente en ambos “exones”), que si tuviera un ORF funcional en estos dos individuos tendría un tamaño de 505 pb.

Para los *primers* 14 se usó una temperatura de hibridación de 55,4°C en la amplificación del ADNg con NZYTaQ II (mostrada en la figura 19B), aunque posteriormente se hizo un gradiente de temperaturas en el que se determinó 64,1°C como la óptima (en la que desaparecían las bandas inespecíficas). En el gel de la figura 19B se observan bandas ligeramente por encima de las 200 pb para todas las muestras, que *a priori* parece coincidir con el producto esperado de 196 pb. También había bandas inespecíficas bastante prominentes en algunas de las muestras (macho 1, hembra 1 y hembra 3). Aunque la PCR *in silico* resultó en 27 posibles productos adicionales entre 109 y 2.864 pb con los *primers* 14 Fw-Fw, ninguno de los tamaños se corresponde con las bandas observadas en este gel. Se desconoce, por tanto, su origen; pero se podría descartar una contaminación de los reactivos como agente causante, dado que el blanco sólo presenta una banda correspondiente a dímeros de *primers* (por encima de 50 pb). Sí que se podría tratar de una contaminación de las muestras de ADN genómico.

Con respecto a la amplificación de ADNc con estos mismos *primers* 14, en la figura 19A no se observa el producto esperado en ninguna de las muestras (con una consecuente aparición de dímeros de *primers* en todas las demás muestras, al haber mucha menos amplificación) excepto quizás en la muestra 6, donde aparecen dos bandas (una de 200 pb y otra por encima, de unas 250 pb) que recuerdan a las dos bandas de tamaño similar observadas en la muestra H2 de la figura 19B (para ADNg). La primera de estas bandas podría ser el producto buscado; aunque la PCR *in silico* proporciona fragmentos de 223 y 275 pb, se desconoce si la banda ligeramente mayor podría tener esa procedencia. Por último, para los individuos 1 y 4 se distinguen dos bandas de 700 pb que podrían tratarse del producto inespecífico de 684 pb obtenido en la PCR *in silico*.

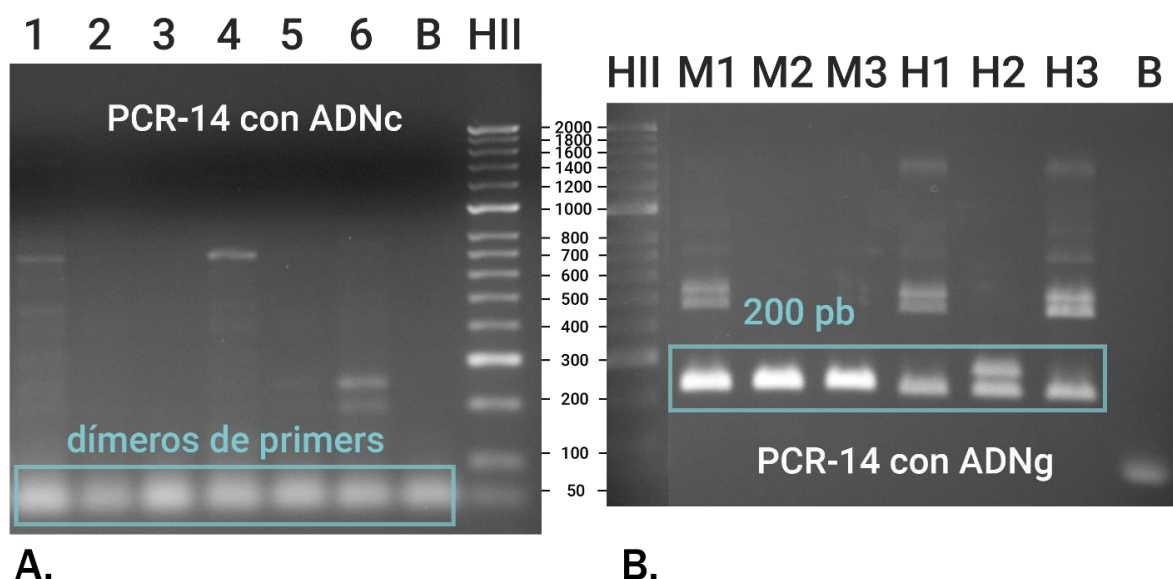


Figura 19. A: amplificación de ADNc con NZYTaQ II y con *primers dmrt1-14*. 1-6: ADNc de macho inmaduro WT, hembra inmadura WT, macho inmaduro F1, hembra inmadura F1, macho maduro F1, hembra madura F1. B: blanco/control negativo. HII: HyperLadder II. **B:** amplificación de ADNg con NZYTaQ II con *primers dmrt1-14*. HII: HyperLadder II. M1-H3: ADNg de tres individuos macho y tres individuos hembra. B: blanco/control negativo.

En general, como se expone en mayor detalle en el apartado de qPCR, los *primers dmrt1-13* y *14* no parecen funcionar correctamente con el ADNc. La segunda pareja sí da resultados para el ADNg, pero el éxito de los *primers 13* para el mismo es incierto debido a la incapacidad de distinguir el producto de lo que aparentan ser dímeros de *primers*.

7.7. Escisión y reamplificación de banda de 1.800 pb de PCR-10

Una vez cortadas las bandas de 1.800 pb de ambos individuos (figura 20A), purificadas y resuspendidas, se hicieron reamplificaciones con las mismas condiciones de PCR para obtener productos con una concentración suficiente para ser secuenciados. Ninguna de los intentos dio el tamaño de banda buscado; en una de las reamplificaciones (figura 20B) se obtuvieron en su lugar bandas por debajo de ~300 y ~250 pb, además de una entre 100 y 200 pb, para todas las muestras (y dímeros de *primers* en las muestras M1' y H2'). Algunas de estas bandas indeseadas estaban presentes en el gel original del que se cortaron las bandas, tanto en las muestras como en los blancos (no mostrados en la imagen), por lo que la contaminación habría tenido origen en los reactivos de PCR. No fue posible, por ende, secuenciar la banda desconocida e inespecífica de 1.800 pb.

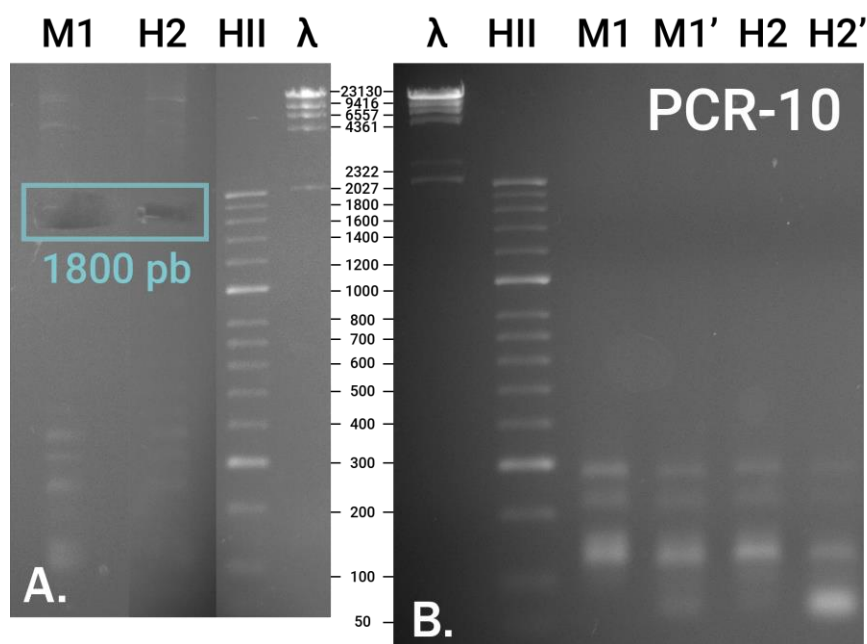


Figura 20. A: amplificación de ADN_g de dos individuos seleccionados con Q5 y con *primers dmrt1-10*, tras cortar las bandas de interés de 1.800 pb. M1: macho 1. H2: hembra 2. HII: HyperLadder II. λ: Lambda DNA/HindIII. **B:** reamplificación de las muestras purificadas correspondientes a las bandas con Q5 y con *primers dmrt1-10*. λ: Lambda DNA/HindIII. HII: HyperLadder II. M1: 1 µL de la muestra de banda para el macho 1. M1': 5 µL de la muestra de banda para el macho 1. H2: 1 µL de la muestra de banda para la hembra 2. H2': 5 µL de la muestra de banda para la hembra 2.

7.8. Análisis de polimorfismos mediante digestión con enzimas de restricción del intrón 3

La tabla 17 a continuación proporciona una lista generada por BioEdit de aquellas enzimas de restricción que cortan menos de 5 veces en la secuencia amplificada por los *primers dmrt1-12* (que con 2.522 pb constituye la totalidad del intrón 3), que *a priori* serían adecuadas para el experimento.

Tabla 17. Lista de enzimas de restricción que cortan menos de 5 veces en el intrón 3 de *dmrt1*.

En azul y en **negrita**, las enzimas finalmente elegidas.

Enzima	Secuencia de reconocimiento	Frecuencia	Posiciones
<i>AflII</i>	C'TTAA_G	1	2149
<i>BseGI</i>	GGATG_nn'	4	213, 724, 2027, 2466
<i>Bsp1407I</i>	T'GTAC_A	2	784, 1961
<i>BsrGI</i>	T'GTAC_A	2	784, 1961
<i>BsuRI</i>	GG'CC	3	1389, 1395, 1478
<i>Cfr13I</i>	G'GnC_C	3	634, 1388, 1477
<i>DraI</i>	TTT'AAA	2	97, 1667
<i>EarI</i>	CTCTTCn'nnn_	1	988
<i>Eco32I</i>	GAT'ATC	1	819
<i>EcoRV</i>	GAT'ATC	1	819
<i>HhaI</i>	G_CG'C	4	1151, 1294, 1810, 1812
<i>Hin6I</i>	G'CG_C	4	1149, 1292, 1808, 1810
<i>HinP1I</i>	G'CG_C	4	1149, 1292, 1808, 1810
<i>HindIII</i>	A'AGCT_T	1	791
<i>PstI</i>	C_TGCA'G	2	1979, 2336
<i>TaqI</i>	T'CG_A	1	1859

Finalmente, se escogieron *DraI* (que produce fragmentos de 97, 1.570 y 952 pb) y *PstI* (resultante en fragmentos de 1.979, 357 y 186 pb) en base a su disponibilidad y a la facilidad de observar los fragmentos en gel de agarosa, por sus tamaños fáciles de diferenciar. El gel resultante de la digestión se muestra en la figura 21: a juzgar por la intensidad y tamaño de bandas, las digestiones parecen haber sido completas ya que no es posible ver bandas de producto PCR sin digerir (que tendrían un mayor tamaño, de 2.522 pb).

Los tamaños de los fragmentos concuerdan con los esperados para ambas enzimas de restricción, aunque son difíciles de observar para las tres primeras muestras (y especialmente la del macho 1) debido a poca cantidad de producto PCR del que se partía (probablemente por una degradación del ADN_g durante la conservación de estas muestras). De cualquier manera, si no hubiera habido digestión en esos sitios de reconocimiento para estas muestras, alguno de los otros dos fragmentos habría migrado menos en el gel (al tener un mayor tamaño) lo cual no ha ocurrido, de forma que se puede afirmar que los seis individuos analizados son monomórficos e idénticos entre sí para las secuencias de reconocimiento de las enzimas de restricción *DraI* y *PstI* en el intrón 3 del gen *dmrt1*. Es importante mencionar que en este caso al hablar del intrón 3 nos referimos a la secuencia entre los exones III y II' (es decir, incluye también al exón no funcional $\Delta I'$ y al intrón 1'; no es el intrón 3 representado en la figura 4 de la introducción), porque $\Delta I'$ no se trata de un exón como tal.

Estas cuatro secuencias de reconocimiento (dos TTAAA y dos CTGCAG) también se encontraban presentes en la secuencia del genoma del lenguado. De hecho, el alineamiento entre las secuencias del intrón 3 obtenidas a partir de diferentes estudios e individuos [35, 44] resultó en una coincidencia del 98,98%, con 10 sitios con gaps y 15 sitios polimórficos, sugiriendo un cierto grado de conservación relativamente alto de este intrón en el lenguado (aunque sería necesario comparar las secuencias de muchos más individuos para confirmarlo). Si tras llevar a cabo otros estudios esto resultara ser cierto, podría ser indicativo de que el intrón presenta una función en la expresión del gen, por ejemplo, por contener secuencias funcionales como genes codificantes de ARN de pequeño tamaño), considerando además que los genes evolutivamente conservados tienden a ganar intrones, no a perderlos (como en este caso) [52].

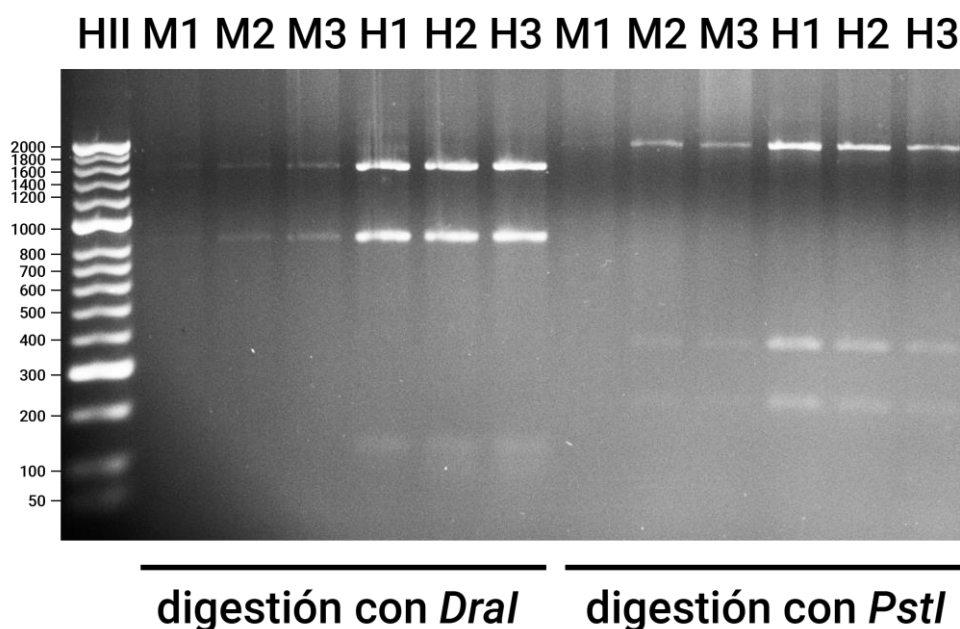


Figura 21. Digestión simple con las enzimas de restricción *DraI* y *PstI* del producto PCR previamente obtenido con los *primers* *dmrt1*-12 y NZYTaqII (intrón 3). HII: HyperLadder II. M1-H3: ADN_g de tres individuos macho y tres individuos hembra.

7.9. Retrotranscripción de ARN a ADNc

Después de la retrotranscripción de las 6 muestras facilitadas de ARN, se midieron las absorbancias por espectrofotometría para determinar la concentración y la pureza de las muestras, presentándose los resultados en la tabla 18. Los valores de concentración de ADNc no se pueden tomar como fiables, ya que la técnica no distingue los dNTPs y *primers* sobrantes (ni del ARN que podría no haberse retrotranscrito) del propio ADNc.

Tabla 18. Resultados de la cuantificación previa del ARN y de la cuantificación y determinación de la pureza de las muestras de ADNc resultantes de la retrotranscripción del ARN. A260/280 para ARN puro: 2. A260/230 para ADN y ARN puro: 2-2,2.

Muestra	Concentración ARN (ng/μL)	Concentración ADNc (ng/μL)	A260/A280	A260/A230
1: Macho inmaduro WT	1438,3	1278,8	1,82	2,27
2: Hembra inmadura WT	832,8	1457,1	1,80	2,13
3: Macho inmaduro F1	361,3	1634,3	1,80	1,83
4: Hembra inmadura F1	118,7	1450,7	1,79	1,99
5: Macho maduro F1	2036,1	1462,6	1,81	2,12
6: Hembra madura F1	427,8	1385	1,80	2,13

7.10. PCR cuantitativa (qPCR) (con *primers dmrt1-13 y 14*)

De la primera prueba de qPCR con los *primers dmrt1-13 y 14*, en la que se hizo un gradiente de temperaturas para dos individuos, se obtuvieron dos picos muy diferenciados en la curva de *melting* de la figura 22: 83,5°C para *primers 13* y 75°C para los *primers 14*, en todas las muestras en ambos casos.

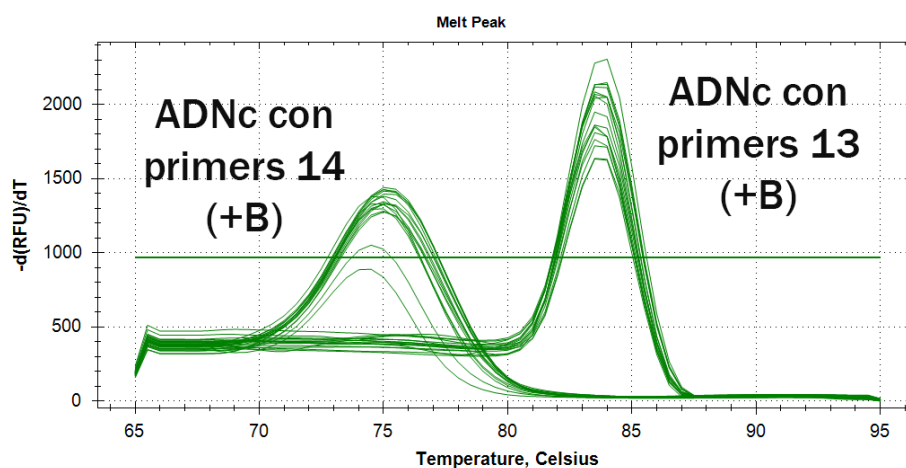


Figura 22. Curva de *melting* para gradiente de qPCR con los *primers dmrt1-13 y 14* y ADNc de dos individuos.

Ya que los blancos (NTC, control sin muestra o *no template control*) también daban señales de fluorescencia altas (en ocasiones, mayores que algunas de las muestras) con una temperatura de *melting* muy similar a la de las muestras, si no idéntica, se pensó que podría haber contaminación. El gel de comprobación de esta qPCR confirmó los resultados: en los NTC había bandas tan intensas como las muestras, independientemente de la temperatura de hibridación, por lo que se optó por usar 60°C en los siguientes experimentos.

Se prepararon 7 reacciones más, en las que se probó una variedad de condiciones para eliminar la contaminación: preparar las placas en distintas condiciones de esterilidad (en llama y en cabina de flujo laminar), utilizar un kit de SYBR Green distinto (NZYSupreme qPCR Green Master Mix), emplear nuevas alícuotas de *primers* (que tuvieron que ser pedidas de nuevo) y probar con distintas concentraciones de *primers*. Después de comprobar las muestras en gel, se siguió observando bandas en los NTC.

Otra de las pruebas fue utilizar los *primers eEf1a1a* modo de control positivo, con 4 muestras de ADNc y 4 muestras de ADNg de distintos individuos y NTC por triplicado, para contrastar los resultados de los NTC con los de los *primers* 13. Del gel de comprobación de esta qPCR (figura 23) se verifica que el causante de estos resultados no es contaminación, sino que es la formación de dímeros de *primers* 13 (en los NTC y al utilizar ADNc como muestra) que no ocurre para *eEf1a1*. De acuerdo con el aspecto y tamaño de las bandas, la amplificación fue exitosa para ambos tipos de muestras con los *primers eEf1a1*, pero sólo tuvo lugar con ADNg en el caso de los *primers* 13 dado que para el ADNc se observan bandas difusas y poco intensas, idénticas a las de los NTC. Las temperaturas de *melting* para todas las muestras con *primers* 13 (incluidos los tres NTC) son de 83,5°C, mientras que para los *primers* control son de 84°C (y nulas para los NTC), poniendo de manifiesto la imposibilidad de distinguir dímeros de *primers* 13 del producto esperado exclusivamente en base a la curva de *melting*. De esto se concluye que los *primers* 13 no son aptos para qPCR, pues forman un alto número de dímeros (incluso a bajas concentraciones de *primers* según lo observado en otros experimentos) que además presentan una temperatura de *melting* idéntica al producto esperado, impidiendo así la obtención de resultados fiables.

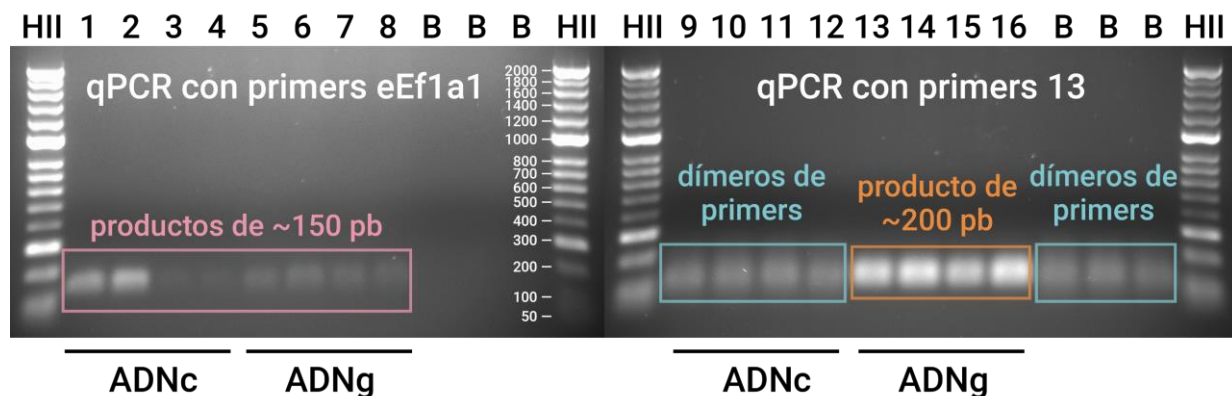


Figura 23. Amplificación por qPCR de ADNg y ADNc de 4 individuos (cada uno) con TB Green Premix Ex Taq II y con *primers eEf1a1* (control de expresión de referencia) y 13. HII: HyperLadder II. 1-4 y 9-12: ADNc de macho inmaduro WT, hembra inmadura WT, macho inmaduro F1 y hembra inmadura F1. 5-8 y 13-16: ADNg de macho 1, macho 2, hembra 1 y hembra 2. B: blanco o NTC.

Para comprobar si las muestras de ADNc verdaderamente no daban los productos esperados, se llevó a cabo una última qPCR en la que se empleó cuatro muestras de ADNc y otras cuatro de ADNg, de diferentes individuos, tanto con los *primers* 13 como los 14.

Junto a los resultados anteriores, la figura 24 parece indicar que ambas parejas de *primers* son incapaces de amplificar el ADNc, pero con los *primers* 14 sí que se obtiene un producto con una temperatura de *melting* bastante distinta, de 87,5°C.

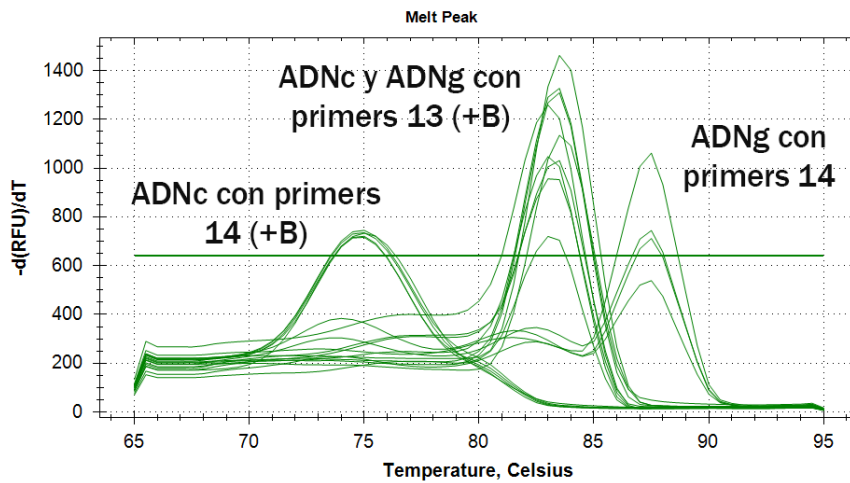


Figura 24. Curva de *melting* para qPCR con los *primers* *dmrt1*-13 y 14, con cuatro muestras de ADNc y cuatro muestras de ADNg.

Por último, se corrieron las muestras en gel de agarosa para verificar el tamaño de los productos obtenidos (figura 25). Confirmando los previos resultados, no se observa amplificación al utilizar ADNc para ninguno de los *primers*, obteniéndose dímeros de *primers* muy similares a los NTC. Una peculiaridad es la aparición de dos bandas distintas para los dímeros de *primers* 13, por encima y por debajo de 100 pb; en geles anteriores se apreciaba una sola banda, más difusa y muy poco definida, lo cual debe estar relacionado con la resolución del gel.

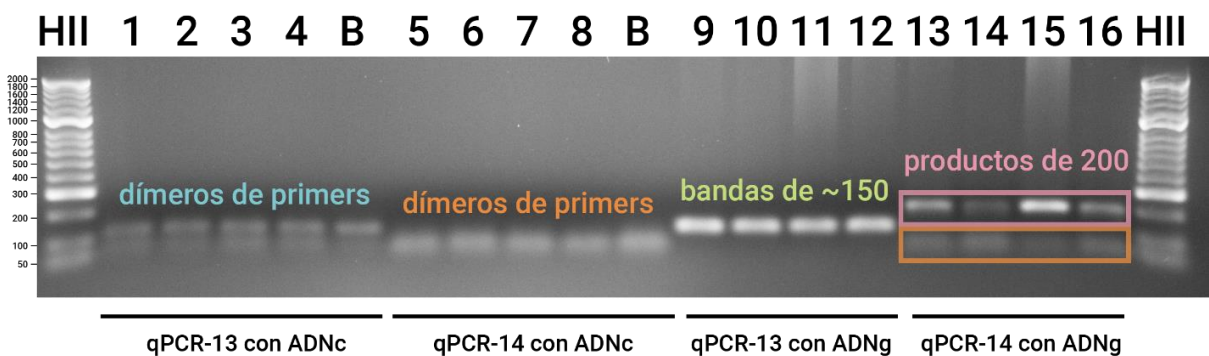


Figura 25. Amplificación por qPCR de ADNg y ADNc de 4 individuos (cada uno) con TB Green Premix Ex Taq II y con *primers* 13 y 14. HII: HyperLadder II. 1-4 y 5-8: ADNc de macho inmaduro WT, hembra inmadura WT, macho inmaduro F1 y hembra inmadura F1. 9-16: ADNg de macho 1, hembra 1, macho 2 y hembra 2. B: blanco o NTC.

Sin embargo, sí se ve amplificación al usar ADNg. Para los *primers* 13 aparecen bandas de unas 150 pb, que aunque no parecen corresponderse con el producto esperado (110 pb) se contempla la posibilidad de que los dímeros de *primers* se estén uniendo al producto y aumentando el tamaño del mismo (lo que explicaría la completa ausencia de dímeros de *primers* en las muestras con ADNg, teniendo en cuenta además que los dímeros y el producto presentan la misma temperatura de *melting*).

Por otro lado, los *primers* 14 dan, además de dímeros de *primers*, una única banda de 200 pb que sí concuerda con el producto esperado. Estos resultados coinciden con los obtenidos durante la amplificación por PCR estándar con estas mismas parejas: no se obtienen productos con ADNc pero sí con ADNg. Ya que sí se consiguió amplificar correctamente el ADNc con los *primers* *eEf1a1* (de referencia), esto no señala a la calidad de las muestras de ADNc en sí como causa del problema. Una posible explicación es que el gen no tenga expresión en el tejido, aunque al tratarse de gónadas es improbable, puesto que es el lugar de actuación. Además, otro estudio sí fue capaz de observar expresión en este tejido [35]. Se podría achacar la falta de amplificación a la temperatura de hibridación empleada (de 60°C); sin embargo, se hicieron dos gradientes de temperatura de 55 a 65°C para ambos *primers*, que fueron comprobados en gel, en los que no se hallaron diferencias. Por tanto, se desconoce por completo el origen de estos resultados negativos para ADNc, tanto con NZYtaq II como con TB Green Premix Ex Taq II. De todas maneras, incluso si se hubiera conseguido amplificación del ADNc los *primers* no podrían ser utilizados en qPCR debido a la gran cantidad de dímeros de *primers* que forman.

7.11. Alineamiento de la secuencia del gen *dmrt1* con el genoma

La versión local de blastn (BLAST+) arrojó una larga lista de resultados de alineamientos de la secuencia del gen (problema) con distintas partes del genoma del lenguado (secuencias sujeto), en la que se obtiene una región relativamente continua (con algunos huecos entre alineamientos consecutivos) de elevada identidad entre la secuencia del gen y el genoma, con una cobertura de 30.638 pb (del 100%), que es el tamaño del gen (excluyendo las regiones 5' y 3'UTR). Estos huecos existentes podrían explicarse por la existencia de regiones de baja complejidad (repeticiones de mononucleótidos y repeticiones cortas) que, por defecto, el algoritmo no considera en los alineamientos, de forma que se producen interrupciones que dan lugar a los huecos observados.

Por su parte, de la versión online de blastn destaca la primera secuencia, en el cromosoma 5 (LG5), que produce alineamientos con una cobertura de la secuencia del gen del 100%, un valor E de 0, una alta puntuación de alineamiento y un porcentaje de identidad del 98,55%, como se ve en la figura 26. La anotación de estas secuencias sujeto para el primer resultado es *doublesex-and mab-3-related transcription factor 1-like* y el número de acceso coincide con el obtenido en BLAST+, por lo que se debe tratar de la misma serie de alineamientos.

	Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/>	Solea senegalensis isolate Sse05_10M linkage group LG5. whole genome shotgun s...	52065	2.390e+05	100%	0.0	98.55%	27692037	CM032858.1
<input type="checkbox"/>	Solea senegalensis isolate Sse05_10M linkage group LG5. whole genome shotgun s...	52065	2.390e+05	100%	0.0	98.55%	27692037	JAGKHQ010000017.1
<input type="checkbox"/>	Solea senegalensis isolate Sse05_10M linkage group LG3. whole genome shotgun s...	504	1.709e+05	16%	2e-139	81.66%	33319822	CM032856.1
<input type="checkbox"/>	Solea senegalensis isolate Sse05_10M linkage group LG3. whole genome shotgun s...	504	1.709e+05	16%	3e-139	81.66%	33319822	JAGKHQ010000015.1
<input type="checkbox"/>	Solea senegalensis isolate Sse05_10M linkage group LG8. whole genome shotgun s...	487	1.893e+05	16%	6e-134	81.91%	27361452	CM032861.1
<input type="checkbox"/>	Solea senegalensis isolate Sse05_10M linkage group LG8. whole genome shotgun s...	487	1.893e+05	16%	7e-134	81.91%	27361452	JAGKHQ010000020.1

Figura 26. Alineamientos más significativos dados por blastn (versión online) entre la secuencia de *dmrt1* previamente descrita en el lenguado y el genoma de un lenguado macho.

La diferencia entre los resultados de BLAST local y online es mínima y probablemente se pueda atribuir a distintos parámetros por defecto en los algoritmos. En ambos casos, existe una secuencia del genoma de referencia (la línea completamente roja de la figura 27, con una alta puntuación de alineamiento) que produce alineamientos significativos que cubren la secuencia conocida del gen *dmrt1*. Por tanto, aunque la coincidencia no es del 100%, es lo suficientemente alta como para afirmar que la secuencia del gen con la que se trabajó a lo largo de este trabajo queda confirmada frente a este genoma ensamblado a partir de un macho.

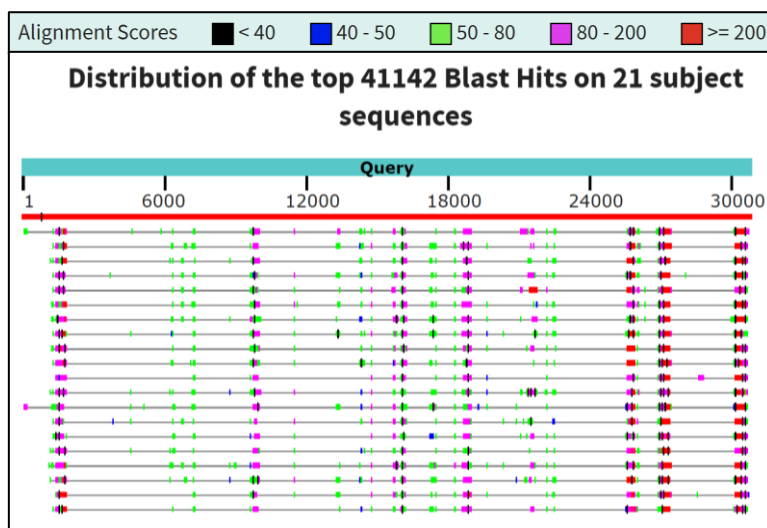


Figura 27. Distribución de los alineamientos más significativos dados por blastn (versión online) entre la secuencia de *dmrt1* previamente descrita en el lenguado y el genoma de un lenguado macho.

Fuente: BLAST.

Por último, se debe hacer notar la gran cantidad de alineamientos obtenidos fuera del gen de interés, mostrados en las 20 secuencias representadas en la figura 27, por debajo de la secuencia roja (que se corresponde al alineamiento del gen consigo mismo). Cubren un pequeño porcentaje de la secuencia problema (entre un 12-17%); no obstante, tienen valores de E (*E-values*) muy bajos (lo que implicaría que no es probable que los alineamientos se deban al azar, haciéndolos bastante significativos) y coincidencias relativamente altas (de alrededor del 80%). Estas secuencias sujeto se encuentran en todos los otros cromosomas (de LG1 a LG21, incluyendo LG5 que es donde está ubicado el gen *dmrt1*). Aunque algunas de estas secuencias se encuentran anotadas (*extracellular calcium-sensing receptor-like*, *lysophosphatidic acid receptor 6-like*, *signal peptide*, *CUB and EGF-like domain-containing protein 1 isoform X1*, *doublesex-and mab-3-related transcription factor a1* y *CUB and EGF-like domain-containing protein 1 isoform X1*, entre muchos otros), otras no lo están y se encuentran entre otros genes que sí están anotados (por ejemplo, *6465 bp at 5' side: adm2-like* y *29132 bp at 3' side: glycosyltransferase 8 domain-containing 2*). Hay que mencionar que los alineamientos se dan casi en la totalidad de la secuencia del gen (como se ve en la figura 27), es decir, tanto intrones como exones presentan homología con otras secuencias del genoma del lenguado. Esto parece estar de acuerdo con los eventos de duplicación del genoma que se han producido a lo largo de la evolución de los teleósteos [51] y podría explicar la abundante amplificación inespecífica observada durante la realización de este trabajo con los *primers* diseñados, como se comentó en uno de los apartados anteriores.

8. CONCLUSIONES

1. La secuencia del gen *dmrt1* previamente obtenida para el lenguado senegalés es prácticamente idéntica a la encontrada en el genoma recién publicado de un macho.
2. No fue posible amplificar y secuenciar las regiones intrónicas del gen *dmrt1* con las parejas de *primers dmrt1*-10, 11, 15 y 16, por no haberse podido obtener los productos PCR esperados (y en cambio, dándose una gran cantidad de amplificación inespecífica).
3. Se encontró un alto grado de homología entre ciertas partes de la secuencia del gen *dmrt1* y otras secuencias fuera del gen repartidas a lo largo del genoma del lenguado.
4. Los seis individuos analizados son monomórficos para las cuatro secuencias de reconocimiento de las enzimas de restricción *DraI* y *PstI* en el intrón 3 del gen *dmrt1*.
5. Las parejas de *primers dmrt1*-13 y 14 no son aptas para realizar PCR cuantitativa por una abundante formación de dímeros de *primers*. No dieron el producto esperado al utilizar ADN complementario en ninguna de las condiciones probadas, pero sí con ADN genómico. Por todo ello, no se pudo determinar la eficiencia de los *primers* ni cuantificar la expresión del gen *dmrt1*.

9. PERSPECTIVAS DE FUTURO

Siguiendo con el objetivo inicial del trabajo, sería recomendable diseñar nuevas parejas de *primers* tanto para el gen completo como para los exones a cuantificar, tratando de que sean lo más específicos posibles (comprobando los posibles productos con la secuencia del genoma completo) y que formen el menor número de dímeros y horquillas. Para reducir la hibridación inespecífica se podría realizar PCR *touchdown* o *hot start* en lugar de PCR estándar. En la primera, se empieza utilizando altas temperaturas de hibridación que van reduciéndose por incrementos cada ciclo. De esta forma, los *primers* hibridan específicamente a la mayor temperatura dando como primera secuencia amplificada el producto deseado. A medida que baja la temperatura de hibridación, los fragmentos deseados siguen amplificándose y disminuye la probabilidad de que se obtengan productos inespecíficos, aumentando además la eficiencia de la reacción hacia el final del programa al aumentar la estabilidad del dúplex *primer*-molde con la reducción de la temperatura. Por otra parte, la PCR *hot start* reduce la cantidad de productos inespecíficos y dímeros de *primers* mediante la inactivación de la polimerasa hasta ser activada por calor, o bien mediante una adición tardía de la misma.

En términos más generales, para verificar la hipótesis de *dmrt1* como candidato a gen maestro regulador de la determinación sexual en *S. senegalensis* sería imprescindible analizar las secuencias del gen en múltiples individuos para buscar diferencias que puedan señalar su posible función; por ejemplo, en el caso de la medaka, la copia duplicada del gen sólo se encuentra en el cromosoma Y de los machos [18]. Con el mismo propósito se debería tratar de identificar la región determinante del sexo (SDR o SD, *sex-determining region*) a través de mapas de marcadores y de QTL (loci de caracteres cuantitativos, o *quantitative trait loci*) asociados al sexo, pues el gen maestro debería encontrarse en esta región, lo que ayudaría a delimitar la búsqueda. Para esto se debe seguir realizando estudios de asociación del genoma completo (GWAS, *genome-wide association study*) con más individuos, como ya se viene haciendo [44], para identificar marcadores asociados con el sexo que además de permitir dilucidar la SDR ayudarían a la tarea de identificación del sexo de los peces y a poner en marcha programas de cría selectiva para mejora genética.

Asimismo, se podrían realizar estudios de expresión génica de *dmrt1* por qPCR en diferentes etapas de la vida del lenguado, permitiendo así relacionarla con el proceso de maduración de las gónadas en el tiempo. También resultaría interesante emplear muestras de individuos cultivados bajo distintas condiciones de fotoperiodos, nutrición y, especialmente, de temperatura del agua, puesto que hay evidencia de la influencia de los termociclos diarios en el desarrollo de las gónadas y en la relación macho/hembra en esta especie [27], fenómeno que *dmrt1* podría estar regulando.

Una vez hallado el maestro regulador de la determinación sexual, sea *dmrt1* u otro, todavía se tendría que validar su función específica de forma experimental, además de otros genes relacionados, para obtener una visión más general de los procesos de determinación y diferenciación sexual que ayuden a desarrollar técnicas para la acuicultura del lenguado senegalés, tales como la obtención de poblaciones de solo hembras por ginogénesis artificial.

10. BIBLIOGRAFÍA

- [1] Cañavate, P. (2013). Diversificación de especies en la piscicultura marina española. Editorial Instituto Español de Oceanografía. *Cultivo de soleidos: el lenguado senegalés*. 249-283.
- [2] Rodríguez Villanueva, J.L. y Peleteiro Alonso, J.B. (2014). Cultivo del lenguado senegalés (*Solea senegalensis*). Editorial Fundación Observatorio Español de Acuicultura.
- [3] Morais, S., Aragão, C., Cabrita, E., Conceição, L.E.C., Constenla, M., Costas, B., et al. (2016). New developments and biological insights into the farming of *Solea senegalensis* reinforcing its aquaculture potential. *Reviews in Aquaculture*. 8 (3), 227-263.
- [4] Arias, A.M. y de la Torre, M. (2019). *Solea senegalensis* (nombre científico) / Ictiobase.es. Recuperado de http://www.ictiobase.es/nombre_cientifico.php?nc=183 el 26 de mayo de 2021.
- [5] López, J.M.A. y Reiriz, M. del C.S. (2005). Atlas histológico del lenguado senegalés, «*Solea senegalensis*» (Kaup, 1858). Editorial Consejo Superior de Investigaciones Científicas.
- [6] Viñas, J., Asensio, E., Cañavate, J.P., y Piferrer, F. (2013). Gonadal sex differentiation in the Senegalese sole (*Solea senegalensis*) and first data on the experimental manipulation of its sex ratios. *Aquaculture*. 384-387, 74-81.
- [7] Muñoz-Cueto, J.A., Mañanós, E.L., y Sánchez-Vázquez, F.J. (2019). The Biology of Sole. Editorial CRC Press.
- [8] Cañavate, J.P. (2005). Opciones del lenguado senegalés *Solea senegalensis* Kaup, 1858 para diversificar la acuicultura marina. *Boletín del Instituto Español de Oceanografía*. 21 (1-4), 147-154.
- [9] APROMAR (2020). La Acuicultura en España 2020. 55.
- [10] Duncan, N., Carazo, I., Norambuena, F., Estevez, A., Chereguini, O., y Mañanos, E. (2014). Reproductive behaviour of the Senegalese sole (*Solea senegalensis*): a review with indications of possible behavioural solutions to the G1 reproductive dysfunction [Acta de conferencia]. Aquaculture Europe, San Sebastián, España.
- [11] Chereguini, O., Rasines, I., Anguis, V., Martín, I., Rodríguez, C., Guzmán, J.M., et al. (2007). Primeras fecundaciones artificiales en lenguado senegalés cultivado (Generación F1) [Acta de conferencia]. XI Congreso Nacional de Acuicultura, Xunta de Galicia. Vigo, España. 1435-1438.
- [12] Rasines, I., Gómez, M., Martín, I., Rodríguez, C., Mañanós, E., y Chereguini, O. (2012). Artificial fertilization of Senegalese sole (*Solea senegalensis*): Hormone therapy administration methods, timing of ovulation and viability of eggs retained in the ovarian cavity. *Aquaculture*. 326-329, 129-135.
- [13] Mei, J. y Gui, J.F. (2015). Genetic basis and biotechnological manipulation of sexual dimorphism and sex determination in fish. *Science China Life Sciences*. 58 (2), 124-136.
- [14] Brykov, V.A. (2014). Mechanisms of sex determination in fish: Evolutionary and practical aspects. *Russian Journal of Marine Biology*. 40 (6), 407-417.
- [15] Heule, C., Salzburger, W., y Böhne, A. (2014). Genetics of sexual development: An evolutionary playground for fish. *Genetics*. 196 (3), 579-591.
- [16] Trukhina, A. V., Lukina, N.A., Wackerow-Kouzova, N.D., y Smirnov, A.F. (2013). The variety of vertebrate mechanisms of sex determination. *BioMed Research International*. 2013, 1-8.
- [17] Piferrer, F., Ribas, L., y Díaz, N. (2012). Genomic Approaches to Study Genetic and Environmental Influences on Fish Sex Determination and Differentiation. *Marine Biotechnology*. 14 (5), 591-604.
- [18] Nanda, I., Kondo, M., Hornung, U., Asakawa, S., Winkler, C., Shimizu, A., et al. (2002). A duplicated copy of *DMRT1* in the sex-determining region of the Y chromosome of the medaka, *Oryzias latipes*. *Proceedings of the National Academy of Sciences of the United States of America*. 99 (18), 11778-11783.
- [19] Vega, L., Díaz, E., Cross, I., y Rebordinos, L. (2002). Caracterizaciones citogenética e isoenzimática del lenguado *Solea senegalensis* Kaup, 1858. *Boletín del Instituto Español de Oceanografía*. 18 (1-4), 245-250.
- [20] Úbeda, F., Patten, M.M., y Wild, G. (2014). On the origin of sex chromosomes from meiotic drive. *Proceedings of the Royal Society B: Biological Sciences*. 282(1798): 20141932.
- [21] Portela-Bens, S., Merlo, M.A., Rodríguez, M.E., Cross, I., Manchado, M., Kosyakova, N., et al. (2017). Integrated gene mapping and syntenic studies give insights into the evolution of a sex proto-chromosome in *Solea senegalensis*. *Chromosoma*. 126 (2), 261-277.
- [22] Merlo, M.A., Iziga, R., Portela-Bens, S., Cross, I., Kosyakova, N., Liehr, T., et al. (2017). Analysis of the histone cluster in Senegalese sole (*Solea senegalensis*): Evidence for a divergent evolution of two canonical histone clusters. *Genome*. 60 (5), 441-453.
- [23] García-Angulo, A., Merlo, M.A., Portela-Bens, S., Rodríguez, M.E., García, E., Al-Rikabi, A., et al. (2018). Evidence for a Robertsonian fusion in *Solea senegalensis* (Kaup, 1858) revealed by zoo-FISH and comparative genome analysis. *BMC Genomics*. 19 (1), 1-11.
- [24] Rodríguez, M.E., Molina, B., Merlo, M.A., Arias-Pérez, A., Portela-Bens, S., García-Angulo, A., et al. (2019). Evolution of the proto sex-chromosome in *Solea senegalensis*. *International Journal of Molecular Sciences*. 20 (20), 5111.
- [25] Merlo, M.A., Portela-Bens, S., Rodríguez, M.E., García-Angulo, A., Cross, I., Arias-Pérez, A., et al. (2021). A comprehensive integrated genetic map of the complete karyotype of *Solea senegalensis* (Kaup 1858). *Genes*. 12 (1), 1-12.

- [26] Pardo, B.G., Bouza, C., Castro, J., Martínez, P., y Sánchez, L. (2001). Localization of ribosomal genes in Pleuronectiformes using Ag-, CMA3-banding and in situ hybridization. *Heredity*. 86 (5), 531-536.
- [27] Blanco-Vives, B., Vera, L.M., Ramos, J., Bayarri, M.J., Mañanós, E., y Sánchez-Vázquez, F.J. (2011). Exposure of larvae to daily thermocycles affects gonad development, sex ratio, and sexual steroids in *Solea senegalensis*, Kaup. *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology*. 315 A (3), 162-169.
- [28] Dong, J., Li, J., Hu, J., Sun, C., Tian, Y., Li, W., et al. (2020). Comparative Genomics Studies on the dmrt Gene Family in Fish. *Frontiers in Genetics*. 11 (noviembre), 1-17.
- [29] Herpin, A. y Scharl, M. (2011). *Dmrt1* genes at the crossroads: A widespread and central class of sexual development factors in fish. *FEBS Journal*. 278 (7), 1010-1019.
- [30] Matsuda, M., Nagahama, Y., Shinomiya, A., Sato, T., Matsuda, C., Kobayashi, T., et al. (2002). DMY is a Y-specific DM-domain gene required for male development in the medaka fish. *Nature*. 417 (6888), 559-563.
- [31] Otake, H., Shinomiya, A., Matsuda, M., Hamaguchi, S., y Sakaizumi, M. (2006). Wild-derived XY sex-reversal mutants in the medaka, *Oryzias latipes*. *Genetics*. 173 (4), 2083-2090.
- [32] Otake, H., Masuyama, H., Mashima, Y., Shinomiya, A., Myosho, T., Nagahama, Y., et al. (2010). Heritable artificial sex chromosomes in the medaka, *Oryzias latipes*. *Heredity*. 105 (3), 247-256.
- [33] Chen, S., Zhang, G., Shao, C., Huang, Q., Liu, G., Zhang, P., et al. (2014). Whole-genome sequence of a flatfish provides insights into ZW sex chromosome evolution and adaptation to a benthic lifestyle. *Nature Genetics*. 46 (3), 253-260.
- [34] Molina-Luzón, M.J., Hermida, M., Navajas-Pérez, R., Robles, F., Navas, J.I., Ruiz-Rejón, C., et al. (2015). First Haploid Genetic Map Based on Microsatellite Markers in Senegalese Sole (*Solea senegalensis*, Kaup 1858). *Marine Biotechnology*. 17 (1), 8-22.
- [35] Cross, I., García, E., Rodríguez, M.E., Arias-Pérez, A., Portela-Bens, S., Merlo, M.A., et al. (2020). The genomic structure of the highly conserved *dmrt1* gene in *Solea senegalensis* (Kaup, 1868) shows an unexpected intragenic duplication. *PLoS ONE*. 15 (11 de noviembre), 1-21.
- [36] Keren, H., Lev-Maor, G., y Ast, G. (2010). Alternative splicing and evolution: Diversification, exon definition and function. *Nature Reviews Genetics*. 11 (5), 345-355.
- [37] Gao, X. y Lynch, M. (2009). Ubiquitous internal gene duplication and intron creation in eukaryotes. *Proceedings of the National Academy of Sciences of the United States of America*. 106 (49), 20818-20823.
- [38] Koressaar, T. y Remm, M. (2007). Enhancements and modifications of primer design program Primer3. *Bioinformatics*. 23 (10), 1289-1291.
- [39] Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., et al. (2012). Primer3-new capabilities and interfaces. *Nucleic Acids Research*. 40 (15), 1-12.
- [40] Thornton, B. y Basu, C. (2011). Real-time PCR (qPCR) primer design using free online software. *Biochemistry and Molecular Biology Education*. 39 (2), 145-154.
- [41] Chuang, L.Y., Cheng, Y.H., y Yang, C.H. (2013). Specific primer design for the polymerase chain reaction. *Biotechnology Letters*. 35 (10), 1541-1549.
- [42] Orozco-Ugarriza, M.E., Franco-Anaya, P.A., y Olivo-Martínez, Y. (2016). Validación *in silico* de oligonucleótidos-primers para la detección específica de *Salmonella* spp. mediante reacción en cadena de la polimerasa. *Revista de Investigación Agropecuaria y Desarrollo Sostenible*. 1 (1), 42-50.
- [43] Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research*. 31 (13), 3406-3415.
- [44] Guerrero-Cózar, I., Gomez-Garrido, J., Berbel, C., Martinez-Blanch, J.F., Alioto, T., Claros, M.G., et al. (2021). Chromosome anchoring in Senegalese sole (*Solea senegalensis*) reveals sex-associated markers and genome rearrangements in flatfish. *Scientific Reports*. 11 (1), 1-16.
- [45] Slater, G. *Ipcress Manual*. EMBL's European Bioinformatics Institute. Recuperado de <https://www.ebi.ac.uk/about/vertebrate-genomics/software/ipcress-manual> el 15 de agosto de 2021.
- [46] Matlock, B. (2015). Assessment of Nucleic Acid Purity. *Technical Bulletin NanoDrop Spectrophotometers*. 1-2.
- [47] Innis, M.A. y Gelfand, D.H. (1990). *PCR Protocols: A Guide to Methods and Applications*. Optimization of PCRs. Academic Press, San Diego, CA. 3-12.
- [48] O'Neil, S.T. *Command Line BLAST*. A Primer for Computational Biology. Recuperado de <https://www.open.oregonstate.edu/computationalbiology/chapter/command-line-blast/> el 11 de agosto de 2021.
- [49] Altschul, S.F., Gish, W., Miller, W., Myers, E.W., y Lipman, D.J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*. 215 (3), 403-410.
- [50] Kalendar, R., Khassenov, B., Ramankulov, Y., Samuilova, O., y Ivanov, K.I. (2017). FastPCR: An *in silico* tool for fast primer and probe design and advanced sequence analysis. *Genomics*. 109 (3-4), 312-319.
- [51] Volff, J.N. (2005). Genome evolution and biodiversity in teleost fish. *Heredity*. 94 (3), 280-294.
- [52] Carmel, L., Rogozin, I.B., Wolf, Y.I., y Koonin, E. V. (2007). Evolutionarily conserved genes preferentially accumulate introns. *Genome Research*. 17 (7), 1045-1050.

11. AGRADECIMIENTOS

- A mis tutores Alberto Arias y Laureana Rebordinos por guiarme a lo largo del trabajo.
- A Ismael Cross por el fichero de anotación del gen *dmrt1*.
- A Marco Mendizábal por una muestra de ADNg usada en las pruebas iniciales de los *primers*.
- A Javier Aguilar por las seis muestras de ARNm usadas en los experimentos de qPCR.
- A Silvia Portela por los *primers* *eEf1a1*.
- Al resto de miembros del laboratorio por su inestimable ayuda.

12. ANEXOS

12.1. Alineamiento de *dmrt1* con el genoma por BLAST+ (versión local)

Anexo 1. Alineamientos más significativos correspondientes con la posición de *dmrt1* en el genoma, dados por BLAST+ (versión local) entre la secuencia de *dmrt1* previamente descrita en el lenguado y el genoma de un lenguado macho.

En azul y en negrita, la serie de alineamientos que cubre toda la secuencia del gen con un orden lógico. **Saccver**:

versión del código de acceso de la secuencia sujeto. **Pident**: porcentaje de identidad. **Length**: longitud del

alineamiento. **Mismatch**: número de emparejamientos incorrectos. **Gapopen**: número de huecos en la secuencia.

Qstart y **qend**: posición de la secuencia problema en la que comienza y termina el alineamiento. **Sstart** y **send**: posición

de la secuencia sujeto en la que comienza y termina el alineamiento. **Evalue**: valor esperado o valor E, indica la

probabilidad de que el alineamiento se haya dado al azar, por lo que cuanto menor sea su valor más significativo será

el alineamiento. **Bitscore**: a mayor valor, mayor similitud entre las secuencias.

saccver	pident	length	mismatch	gapopen	qstart	qend	sstart	send	evalue	bitscore
CM032858.1	99,483	774	4	0	1	774	15189255	15190028	0,0	1408
CM032858.1	97,152	632	11	1	143	774	15194430	15195054	0,0	1061
CM032858.1	99,116	8370	48	10	771	9134	15190136	15198485	0,0	15023
CM032858.1	99,221	642	5	0	771	1412	15195162	15195803	0,0	1158
CM032858.1	95,009	1122	26	7	2427	3518	15196635	15197756	0,0	1735
CM032858.1	97,683	1381	15	3	5082	6455	15189397	15190767	0,0	2357
CM032858.1	95,336	1115	23	11	7286	8400	15191782	15192867	0,0	1744
CM032858.1	91,291	643	18	4	8659	9263	15193217	15193859	0,0	843
CM032858.1	97,938	4511	19	14	9164	13641	15198484	15202953	0,0	7747
CM032858.1	99,191	4823	25	6	13670	18485	15202950	15207765	0,0	8678
CM032858.1	98,964	12158	71	16	18489	30638	15207729	15219839	0,0	21702

12.2. Alineamiento de *dmrt1* con el genoma por blastn (versión online)

Anexo 2. Alineamientos más significativos fuera de *dmrt1* dados por blastn (versión online) entre la secuencia de

dmrt1 previamente descrita en el lenguado y el genoma de un lenguado macho. **Accession**: versión del código de

acceso de la secuencia sujeto. **Description**: título de la secuencia sujeto partir del archivo FASTA. **Max score**:

puntuación de alineamiento más alta. **Total score**: puntuación total de todos los alineamientos. **Query cover**: porcentaje

de la secuencia problema cubierto en el alineamiento. **E value**: mejor valor esperado o valor E (el más bajo) de entre

todos los alineamientos. **Perc. Ident.**: máxima porcentaje de identidad de entre todos los alineamientos.

Accession	Description	Max Score	Total Score	Query Cover	E value	Per. Ident
CM032856.1	<i>Solea senegalensis</i> isolate Sse05_10M linkage group LG3, whole genome shotgun sequence	504	170900	16	2E-139	81,66
CM032861.1	<i>Solea senegalensis</i> isolate Sse05_10M linkage group LG8, whole genome shotgun sequence	487	189300	16	6E-134	81,91

CM032863.1	<i>Solea senegalensis</i> isolate Sse05_10M linkage group LG10, whole genome shotgun sequence	462	161000	15	3E-126	79,96
CM032865.1	<i>Solea senegalensis</i> isolate Sse05_10M linkage group LG12, whole genome shotgun sequence	453	141900	16	1E-123	76,61
CM032862.1	<i>Solea senegalensis</i> isolate Sse05_10M linkage group LG9, whole genome shotgun sequence	453	141900	15	1E-123	78,77
CM032859.1	<i>Solea senegalensis</i> isolate Sse05_10M linkage group LG6, whole genome shotgun sequence	427	189900	15	5E-116	78,72
CM032855.1	<i>Solea senegalensis</i> isolate Sse05_10M linkage group LG2, whole genome shotgun sequence	421	206700	17	2E-114	76,61
CM032864.1	<i>Solea senegalensis</i> isolate Sse05_10M linkage group LG11, whole genome shotgun sequence	419	163900	16	8E-114	76,5
CM032869.1	<i>Solea senegalensis</i> isolate Sse05_10M linkage group LG16, whole genome shotgun sequence	418	121000	14	3E-113	77,48
CM032857.1	<i>Solea senegalensis</i> isolate Sse05_10M linkage group LG4, whole genome shotgun sequence	414	178400	14	3E-112	78,73
CM032866.1	<i>Solea senegalensis</i> isolate Sse05_10M linkage group LG13, whole genome shotgun sequence	407	122500	12	5E-110	79,87
CM032854.1	<i>Solea senegalensis</i> isolate Sse05_10M linkage group LG1, whole genome shotgun sequence	391	246000	17	4E-105	80,36
CM032867.1	<i>Solea senegalensis</i> isolate Sse05_10M linkage group LG14, whole genome shotgun sequence	380	162300	15	7E-102	77,63
CM032868.1	<i>Solea senegalensis</i> isolate Sse05_10M linkage group LG15, whole genome shotgun sequence	379	133200	13	7E-102	79,45
CM032870.1	<i>Solea senegalensis</i> isolate Sse05_10M linkage group LG17, whole genome shotgun sequence	369	121700	14	1E-98	76,74
CM032872.1	<i>Solea senegalensis</i> isolate Sse05_10M linkage group LG19, whole genome shotgun sequence	359	134100	13	2E-95	77,66
CM032860.1	<i>Solea senegalensis</i> isolate Sse05_10M linkage group LG7, whole genome shotgun sequence	355	181100	17	3E-94	72,66
CM032873.1	<i>Solea senegalensis</i> isolate Sse05_10M linkage group LG20, whole genome shotgun sequence	352	134100	16	1E-93	79,15
CM032871.1	<i>Solea senegalensis</i> isolate Sse05_10M linkage group LG18, whole genome shotgun sequence	340	118400	13	6E-90	76,26